

211.001
V. 60 no. 1 & 2
March-May 1970

CUK-203123-61-POLY414

ALGUTTA
LIBRARY
UNIVERSITY

The American Economic Review

ARTICLES

- W. FELLNER Trends in the Activities Generating Technological Progress
- R. K. DIWAN About the Growth Path of Firms
- D. K. FOLEY AND M. SIDRAUSKI Portfolio Choice, Investment and Growth
- J. P. NEWHOUSE Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital
- M. L. WACHTER Cyclical Variation in the Interindustry Wage Structure
- J. L. STEIN Monetary Growth Theory in Perspective
- A. H. H. TAN Differential Tariffs, Negative Value-Added and the Theory of Effective Protection
- E. E. ZAJAC A Geometric Treatment of Averch-Johnson's Behavior of the Firm Model
- J. R. HARRIS AND M. P. TODARO Migration, Unemployment and Development: A Two-Sector Analysis
- R. A. HOLMES On the Economic Welfare of Victims of Automobile Accidents
- W. H. L. ANDERSON Production Scheduling, Intermediate Goods, and Labor Productivity
- E. F. FAMA Multiperiod Consumption-Investment Decisions
- 61

COMMUNICATIONS: M. Landsberger; P. Zarembka; M. P. Todaro; L. B. Morse; M. V. Pauly; R. L. Crouch; D. Wrightsman; J. H. Makin; H. Hagemann; L. J. Mercer; D. M. Jaffee; S. Peltzman; S. S. Berry; P. Eckstein; I. Adelman and C. T. Morris; J. C. Weicher; B. M. Fleisher.

MARCH 1970

THE AMERICAN ECONOMIC ASSOCIATION

Founded in 1885

P14414

- Published at George Banta Co., Inc., Menasha, Wisconsin.

- THE AMERICAN ECONOMIC REVIEW, including four quarterly numbers, the *Proceedings* of the annual meetings, the *Directory* and Supplements, is published by the American Economic Association and is sent to all members five times a year, in March, May, June, September and December.

- Membership dues of the Association are \$10.00 a year, \$8.00 of each year's dues being in payment of a year's subscription both to the *Review* and the *Journal of Economic Literature*. Subscriptions by nonmembers are \$10.00 a year. Single copies of the *Review*, \$2.00; of the *Journal*, \$2.00.

- Correspondence relating to the *Paper and Proceedings*, the *Directory*, advertising, permission to quote, business matters, subscriptions, membership and changes of address may be sent to the secretary, Harold F. Williamson, 629 Noyes Street, Evanston, Illinois 60201. To be effective, notice of change of address must reach the secretary by the 1st of the month previous to the month of publication.

- Second-class postage paid at Menasha, Wisconsin. Printed in U.S.A.

Officers

President

WASSILY LEONTIEF
Harvard University

President-Elect

JAMES TOBIN
Yale University

Vice-Presidents

EVSEY D. DOMAR
Massachusetts Institute of Technology
TIBOR SCITOVSKY
Yale University

Secretary-Treasurer and Editor of Proceedings

HAROLD F. WILLIAMSON
Northwestern University

Managing Editor of The American Economic Review

GEORGE H. BORTS
Brown University

Managing Editor of The Journal of Economic Literature

MARK PERLMAN
University of Pittsburgh

Executive Committee

Elected Members of the Executive Committee

OTTO ECKSTEIN
Harvard University
ANTHONY SCOTT
University of British Columbia
MARY JEAN BOWMAN
University of Chicago
CHARLES L. SCHULTZE
Brookings Institution
ROBERT DORFMAN
Harvard University
ARNOLD C. HARBERGER
University of Chicago

Ex Officio Members

KENNETH E. BOULDING
University of Chicago
WILLIAM FELLNER
Yale University

C612-1103123-61-P014414.

THE AMERICAN ECONOMIC REVIEW

March 1970

VOLUME LX, NUMBER 1

GEORGE H. BORTS

Managing Editor

WILMA ST. JOHN

Assistant Editor

Board of Editors

GARY S. BECKER

BARBARA R. BERGMANN

JAGDISH N. BHAGWATI

GREGORY C. CHOW

CARL F. CHRIST

C. E. FERGUSON

IRWIN FRIEND

HARRY G. JOHNSON

ALVIN L. MARTY

MARC NERLOVE

G. WARREN NUTTER

VERNON SMITH

WARREN L. SMITH

LLOYD ULMAN

• Manuscripts and editorial correspondence relating to the regular quarterly issue of this REVIEW should be addressed to George H. Borts, Managing Editor of THE AMERICAN ECONOMIC REVIEW, Brown University, Providence, R.I. 02912. Manuscripts should be submitted in duplicate and in acceptable form. *Style Instructions* for guidance in preparing manuscripts will be provided upon request to the editor.

• No responsibility for the views expressed by authors in this REVIEW is assumed by the editors or the publishers, The American Economic Association.

• Copyright American Economic Association 1970.

Articles

Trends in the Activities Generating Technological Progress *William Felner* 1

About the Growth Path of Firms *Romesh K. Diwan* 30

Portfolio Choice, Investment and Growth *Duncan K. Foley and Miguel Sidrauski* 44

Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital *Joseph P. Newhouse* 64

Cyclical Variation in the Interindustry Wage Structure *Michael L. Wachter* 75

Monetary Growth Theory in Perspective *Jerome L. Stein* 85

Differential Tariffs, Negative Value-Added and the Theory of Effective Protection *Augustine H. H. Tan* 107

A Geometric Treatment of Averch-Johnson's Behavior of the Firm Model *E. E. Zajac* 117

Migration, Unemployment & Development: A Two-Sector Analysis *John R. Harris and Michael P. Todaro* 126

On the Economic Welfare of Victims of Automobile Accidents *R. A. Holmes* 143

Production Scheduling, Intermediate Goods, and Labor Productivity *W. H. Locke Anderson* 153

Multiperiod Consumption-Investment Decisions *Eugene F. Fama* 163

Communications

The Life Cycle Hypothesis: A Reinterpretation and Empirical Test	<i>Michael Landsberger</i>	175
Labor Migration and Urban Unemployment:		
Comment	<i>Paul Zarembka</i>	184
Reply	<i>Michael P. Todaro</i>	187
The 1967 Peruvian Exchange Crisis: A Note	<i>Laurence B. Morse</i>	189
Risk and the Social Rate of Discount	<i>Mark V. Pauly</i>	195
A Nonlinear Test of the Random-Walk Hypothesis	<i>R. L. Crouch</i>	199
IS, LM, and External Equilibrium: A Graphical Analysis	<i>Dwayne Wrightsman</i>	203
Reserve Policies of Central Banks:		
Comment	<i>John H. Makin</i>	209
Reply	<i>Helmut A. Hagemann</i>	211
Maximum Bias in Social Saving Estimates Using Prices	<i>Lloyd J. Mercer</i>	212
The Structure of the Money-Expenditures Relationship:		
Comment	<i>Dwight M. Jaffee</i>	216
Reply	<i>Sam Peltzman</i>	220
An Econometric Model of Development:		
Comment	<i>Sara S. Berry</i>	222
Comment	<i>Peter Eckstein</i>	227
Reply	<i>Irma Adelman and Cynthia Taft Morris</i>	236
The Effect of Income on Delinquency:		
Comment	<i>John C. Weicher</i>	249
Reply	<i>Belton M. Fleisher</i>	257
Errata		258
Notes		259

THE AMERICAN ECONOMIC REVIEW



VOL. LX

MAY, 1970

NUMBER 2

PAPERS AND PROCEEDINGS

OF THE

Eighty-second Annual Meeting

OF THE

AMERICAN ECONOMIC ASSOCIATION

New York, New York, December 28-30, 1969

Edited by HAROLD F. WILLIAMSON, *Secretary of the Association*
and

GERTRUDE TAIT, *Executive Assistant*

Copyright American Economic Association, 1970

PRINTED BY GEORGE BANTA COMPANY, INC.

Publication Office: Curtis Reed Plaza, Menasha, Wisconsin 54952

Executive Office: American Economic Association
629 Noyes Street, Evanston, Illinois 60201

Inquiries and other communications regarding membership, meetings and the general affairs of the Association, as well as orders for publications, should be addressed to Harold F. Williamson, Secretary of the American Economic Association, 629 Noyes Street, Evanston, Illinois 60201.

Changes of address should be sent to the Secretary at the Executive Office at 629 Noyes Street, Evanston, Illinois. To be effective, notice must reach the office by the 1st of the month previous to the month of publication. The Association cannot be responsible for copies lost due to failure to report change of address in time.

Entered at the post office at Menasha, Wisconsin, as second class matter. Acceptance for mailing at special rate of postage under the provisions of Sec. 34-40 Par. (D) provided for in the Act of February 28, 1925, embodied in paragraph 4, section 412, P. L. and R., authorized September 13, 1928.

The *American Economic Review* is published five times a year, in March, May, June, September, and December, and is sent to all members of the American Economic Association as one of the privileges of membership, \$8.00 of the annual membership dues of \$10.00 being in payment of a year's subscription to the publication. Printed in U.S.A.

PRICE \$4.00

TABLE OF CONTENTS

Program of the Eighty-second Annual Meeting.....	vii
--	-----

PAPERS

Richard T. Ely Lecture	
The Economics of Production <i>Nicholas Georgescu-Roegen</i>	1

Investment Behavior

A Model of Nonresidential Construction in the United States <i>Charles W. Bischoff</i>	10
Tax Policy and Investment Expenditures in a Model of General Equilibrium <i>Laurits R. Christensen</i>	18
Capacity, Overtime, and Empirical Production Functions <i>Robert E. Lucas, Jr.</i>	23
Discussion	
<i>Arnold C. Harberger</i>	28
<i>Vernon L. Smith</i>	29
<i>Ralph B. Bristol, Jr.</i>	30

Money within the General Framework of the Economic System

Indicators of Monetary Policy: The Arguments and the Evidence <i>Michael J. Hamburger</i>	32
The Choice of Optimal Intermediate Economic Targets <i>Robert Holbrook and Harold Shapiro</i>	40
Implications for Four Econometric Models for the Indicators Issue <i>Richard Zecher</i>	47
Discussion	
<i>Peter A. Frost</i>	55
<i>Leonall C. Andersen</i>	56
<i>Thomas J. Sargent</i>	57

Programming Models of National Economies

Planning with Facts: The Case of Korea <i>Roger D. Norton</i>	59
Quantized Preferences and Planning by Priorities <i>Thomas Vietorisz</i> ..	65
A Dynamic Five-sector Model for Turkey, 1967-82 <i>Charles R. Blitzer, Hikmet Cetin, and Alan S. Manne</i>	70
An Empirical Test of Interregional Input-Output Models: Estimation of 1963 Japanese Production <i>Karen R. Polenske</i>	76

Industrial Organization: Retrospect and Prospect

Industrial Organization: Past History and Future Problems <i>E. T. Grether</i>	83
Industrial Organization: Price Models and Public Policy <i>Eugene M. Singer</i>	90
Industrial Organization: The Role and Contribution of Econometrics <i>Henry Grabowski and Dennis Mueller</i>	100
Discussion	
<i>Peter Asch</i>	105
<i>Almarin Phillips</i>	106
<i>James W. McKie</i>	107

New Dimensions in Natural Resources

New Dimensions for Transnational Marine Resources <i>Francis T. Christy, Jr.</i>	109
The Economics of Congestion and Pollution: An Integrated View <i>Jerome Rothenberg</i>	114
Foreign Enterprises and Developing Nations in the Raw Materials Industries <i>Raymond Vernon</i>	122
Discussion	
<i>Richard R. Nelson</i>	127
<i>Chandler Morse</i>	128
<i>Marshall Hall</i>	129

Basic Data for Policy and Public Decisions: Technical Aspects

Data, Research, and Government <i>Guy H. Orcutt</i>	132
Microdata, Economic Research, and the Production of Economic Knowledge <i>F. Thomas Juster</i> ...	138
A System of Information Centers for Research and Decision Making <i>Charles C. Holl</i>	149
Discussion	
<i>William F. Sharpe</i>	166
<i>Richard Ruggles</i>	167
<i>Edwin Kuh</i>	167

Planning in Free Enterprise Economies

Planning the Foreign Sector in Latin America <i>Carlos F. Diaz-Alejandro</i>	169
--	-----

From Causation to Decision: Planning as Politics <i>Stephen S. Cohen</i>	180	<i>William H. Branson</i>	297
Planning within the Firm <i>Joseph L. Bower</i>	186	<i>Michael Piore</i>	298
Discussion		Market and Price Mechanism in Socialist Countries	
<i>John Sheahan</i>	195	The Rise and Fall of Economic Reform in Czechoslovakia <i>Oldřich Kyn</i>	300
<i>Jesse W. Markham</i>	196	Economic Reforms in Hungary <i>Richard D. Portes</i>	307
<i>Robert A. Solo</i>	197	Growth Performance of Eastern European Economies and Comparable Western European Countries <i>Bela Balassa and Trent J. Bertrand</i>	314
National Communications Policy		Discussion	
The Single-entity Proposal for International Telecommunications <i>Merton J. Peck</i>	199	<i>John M. Montias</i>	321
Technological Advance and Market Structure in Domestic Telecommunications <i>Leland L. Johnson</i>	204	<i>Branko Horvat</i>	323
Spectrum Allocation without Market <i>Harvey J. Levin</i>	209	<i>Evsey D. Domar</i>	324
Discussion		Economics of Education	
<i>Alfred E. Kahn</i>	219	Demand and Supply in U.S. Higher Education: A Progress Report <i>R. Radner and L. S. Miller</i>	326
<i>Roger G. Noll</i>	220	Income Distribution Effects of Higher Education <i>W. Lee Hansen</i>	335
<i>William H. Meckling</i>	222	Resource Allocation in Higher Education <i>Joseph A. Kershaw and Alex M. Mood</i>	341
<i>William K. Jones</i>	223	Discussion	
Economics of Imperialism		<i>William G. Bowen</i>	347
Modern Imperialism: The View from the Metropolis <i>Richard D. Wolff</i>	225	<i>Allan M. Cartter</i>	348
The Structure of Dependence <i>Theotonio Dos Santos</i>	231	<i>Karl A. Fox</i>	349
Militarism and Imperialism <i>Harry Magdoff</i>	237	<i>W. Allen Wallis</i>	350
Discussion		The Teaching of Economics	
<i>Stephen Hymer</i>	243	A Radical Approach to Economics: Basis for a New Curriculum <i>Richard C. Edwards, Arthur MacEwan, and Staff of Social Sciences 125</i>	352
<i>Victor Perlo</i>	244	Economic Efficiency and the Distribution of Benefits from College Instruction <i>W. Lee Hansen, Allen C. Kelley, and Burton A. Weisbrod</i>	364
<i>Arthur MacEwan</i>	245	A Retrospective Look at Undergraduate Economics <i>Ben W. Lewis</i>	370
Distribution Issues: Trends and Policies		Discussion	
Postwar Changes in the Size Distribution of Income in the U.S. <i>Edward C. Budd</i>	247	<i>Paul M. Sweezy</i>	376
Analyzing the American Income Distribution <i>Lester C. Thurow</i>	261	<i>Phillip Saunders</i>	376
Transfer Approaches to Distribution Policy <i>Robert J. Lampman</i>	270	<i>John R. Coleman</i>	378
Discussion		Invited Doctoral Dissertations	
<i>Michael K. Taussig</i>	280	Stochastic Stability of a General	
<i>T. Paul Schultz</i>	281		
<i>Stephan Michelson</i>	283		
Dynamics of Income Distribution: Poverty and Progress			
Variability of Economic Well-being and Its Determinants <i>James D. Smith and James N. Morgan</i>	286		
Discussion			
<i>Arthur M. Okun</i>	296		

CONTENTS

v

Equilibrium Model <i>E. Roy Weintraub</i>	380	Economics in the Industrial State: Science and Sedative	
Farmer's Response to Price in Underdeveloped Areas: The Nicaraguan Case <i>Pedro Belli</i>	385	Economics as a System of Belief <i>John Kenneth Galbraith</i>	469
Equilibrium Growth and Capital Movements between Open Economies <i>Myron G. Myers</i>	393	Discussion <i>Eli Goldston</i>	479
Applications of Engineering Analysis of Production to Econometric Models of the Firm <i>Robert J. Anderson, Jr.</i>	398	<i>Harold Demsetz</i>	481
Discussion <i>Lionel W. McKenzie</i>	403		
<i>Lowell S. Hardin</i> and <i>G. E. Schuh</i>	403	PROCEEDINGS	
<i>Ronald W. Jones</i>	404	Minutes of the Business Meeting	487
Increasing the Supply of Black Economists		John Bates Clark Award	490
Is Economics Culture-bound? <i>Kenneth E. Boulding</i>	406	Minutes of the Executive Committee Meetings	491
Recommendations from Nashville Conference on Economic Curriculum in Black Colleges <i>Flournoy Coles</i>	412	Reports	
An Affirmative Action Plan for the Economics Profession <i>Cleveland A. Chandler</i>	416	Secretary <i>Harold F. Williamson</i> ...	494
The Di-unital Approach to "Black Economics" <i>Vernon J. Dixon</i> ...	424	Treasurer <i>Harold F. Williamson</i> ...	499
The International Firm and Efficient Economic Allocation		Finance Committee <i>C. Wells Farnham</i>	501
International Trade in Inputs and Outputs <i>Robert E. Baldwin</i>	430	Auditor <i>Arthur Andersen & Co.</i> ...	504
Technology Transfer Through the International Firm <i>Jack Baranson</i>	435	Managing Editor, <i>American Economic Review</i> <i>George H. Borts</i> ...	508
The Efficiency (Contradictions) of Multinational Corporations <i>Stephen Hymer</i>	441	Managing Editor, <i>Journal of Economic Literature</i> <i>Mark Perlman</i> ...	513
Discussion <i>Thomas D. Willett</i>	449	Committee on Economic Education <i>G. L. Bach</i>	515
<i>Stefan H. Robock</i>	450	Visiting Scientists Program in Economics <i>Phillip Saunders</i>	517
<i>Ronald I. McKinnon</i>	451	Joint Committee with Association of American Law Schools <i>Gerald M. Meier</i>	518
New Ideas in Pure Theory		Committee on the Functions of the Annual Meeting <i>Kenneth J. Arrow</i>	519
Problems in the Theory of Markets under Uncertainty <i>Roy Radner</i> ..	454	Representative to the Social Science Research Council <i>Albert Rees</i> ...	520
Discussion <i>Ruth P. Mack</i>	461	Representative to the National Bureau of Economic Research <i>Willard L. Thorp</i>	521
<i>Kenneth J. Arrow</i>	462	Representative to the International Economic Association • <i>Fritz Machlup</i>	523
<i>Trout Rader</i>	463	Economics Institute Policy and Advisory Board <i>Henry Rosovsky</i> ...	524
The Economy of Cities		Representatives to the Grove Park Institute <i>Henry H. Villard</i>	525
Discussion <i>Karl A. Fox</i>	465	Informal Meeting of the Black Caucus and A.E.A. Representatives <i>Harold F. Williamson</i>	527
<i>David S. Mundel</i>	466	Statement of Concern of the Caucus of Black Economists <i>Thaddeus H. Sprallen</i>	528
<i>Sam Bass Warner, Jr.</i>	467	Publications of the American Economic Association	531

PROGRAM OF THE EIGHTY-SECOND ANNUAL MEETING OF THE
AMERICAN ECONOMIC ASSOCIATION

New York, New York, December 28-30, 1969

Saturday, December 27, 1969

12:30 P.M.

Executive Committee Meeting

Sunday, December 28, 1969

8:30 A.M.

Industrial Organization: Retrospect and Prospect

Chairman: JESSE W. MARKHAM, Harvard University

Papers: E. T. GREYHER, University of California, Berkeley; EUGENE SINGER, Economic Consultant; DENNIS C. MUELLER, Cornell University, and HENRY GRABOWSKI, Yale University

Discussants: PETER ASCH, Rutgers University; ALMARIN PHILLIPS, University of Pennsylvania; JAMES MCKIE, Vanderbilt University

Increasing the Supply of Black Economists

Chairman: PHYLLIS A. WALLACE, Metropolitan Applied Research Center

Papers: KENNETH E. BOULDING, University of Colorado; FLOURNOY COLES, Fisk University; CLEVELAND CHANDLER, Morgan State College; VERNON DIXON, Princeton University

Discussant: VIVIAN HENDERSON,¹ Clark College

Programming Models of National Economies

Chairman: RICHARD H. DAY, University of Wisconsin

Papers: ROGER NORTON, International Bank for Reconstruction and Development; THOMAS VICTORISZ, New School for Social Research; CHARLES BLITZER, HIKMET CETIN, and ALAN S. MANNE, Stanford University; KAREN POLENSKE, Harvard University

10:30 A.M.

Market and Price Mechanism in Socialist Countries

Chairman: JAROSLAV VANEK, Cornell University

Papers: OLDRYCH KYN, Prague School of Economics; RICHARD PORTES, Princeton University; BELA BALASSA and TRENT BERTRAND, Johns Hopkins University; JAROSLAV VANEK,² Cornell University

Discussants: JOHN M. MONTIAS, Yale University; BRANKO HORVAT, Yugoslav Institute of Economic Research; BENJAMIN WARD,¹ University of California, Berkeley; EVSEY DOMAR, Massachusetts Institute of Technology

The Theory of Interdependent Behavior

Chairman: MARTIN SHUBIK, Yale University

Papers: RICHARD ZECKHAUSER, Harvard University; OTTO A. DAVIS, Carnegie-Mellon University, and ANDREW WHINSTON, Purdue University; MARTIN SHUBIK, Yale University

Discussants: JEROME ROTHEBERG, Massachusetts Institute of Technology; ROBERT WILSON, Stanford University; LESTER TELSER, University of Chicago

Money within the General Framework of the Economic System

Chairman: ALLAN H. MELTZER, Carnegie-Mellon University

Papers: MICHAEL HAMBURGER, Federal Reserve Bank of New York; HAROLD SHAPIRO and ROBERT S. HOLBROOK, University of Michigan; RICHARD ZECHER, University of Chicago

Discussants: PETER FROST, University of Washington; LEONALL ANDERSEN, Federal Reserve Bank of St. Louis; THOMAS SARGENT, U.S. Department of Defense

The Grants Economy as Social Process (Joint session with the Association for the Study of the Grants Economy and the Public Choice Society)

Chairman: KENNETH E. BOULDING, University of Colorado

Papers: MANCUR OLSON, University of Maryland; MARTIN PFAFF and ANITA PFAFF, Wayne State University; ROBERT J. LAMPMAN, University of Wisconsin; JAMES M. BUCHANAN, Virginia Polytechnic Institute; CHARLES J. GOETZ, Virginia Polytechnic Institute

2:00 P.M.

Economics in the Industrial State: Science and Sedative

Chairman: JUANITA KREPS, Duke University

Paper: JOHN K. GALBRAITH, Harvard University

¹ No manuscript received.

² Printed in the December, 1969, *A.E.R.*

Discussants: ELI GOLDSTON, Eastern Gas and Fuel Associates; HAROLD DEMSETZ, University of Chicago

Transportation and Public Utilities (Joint session with the Transportation and Public Utilities Group)

Chairman: WILLIAM VICKREY, Columbia University

Papers: WILLIAM H. MELODY, Federal Communications Commission; MAHLON STRASZHEIM, Harvard University; C. D. FOSTER, Ministry of Transport, London

Discussants: JOSEPH R. ROSE, University of Pennsylvania; JULIUS MALDUTIS, Trans World Airlines; AARON GELLMAN, Budd Company, Philadelphia

3:30 P.M.

Planning in Free Enterprise Economies

Chairman: RAYMOND VERNON, Harvard University

Papers: CARLOS F. DÍAZ-ALEJANDRO, University of Minnesota; STEPHEN COHEN, University of California, Berkeley; JOSEPH BOWER, Harvard University

Discussants: JOHN SHRAHAN, Williams College; JESSE W. MARKHAM, Harvard University; ROBERT SOLO, Michigan State University

Externalities and Public Policy (Joint session with the Econometric Society)

Chairman: JACK HIRSCHLEIFER, University of California, Los Angeles

Paper: KENNETH J. ARROW, Harvard University

Discussants: OTTO A. DAVIS, Carnegie-Mellon University; MANCUR OLSON, University of Maryland; PETER DIAMOND, Massachusetts Institute of Technology

Government and the Grants Economy (Joint session with the Association for the Study of the Grants Economy and the Public Choice Society)

Chairman: GORDON TULLOCK, Virginia Polytechnic Institute

Papers: MICHAEL K. TAUSSIG, Rutgers University; ROY W. BAHL, International Monetary Fund, and JEREMY J. WARFORD, Brookings Institution; RONALD TEEPLES, Claremont College; RALPH NELSON, Queens College; JANOS HORVATH, Butler University

National Communications Policy

Chairman: KERMIT GORDON, Brookings Institution

Papers: MERTON J. PECK, Brookings Institution and Yale University; LELAND JOHNSON, RAND Corporation; HARVEY J. LEVIN, Hofstra University

Discussants: EUGENE V. ROSTOW, Yale University; ALFRED E. KAHN, Cornell University; ROGER G. NOLL, California Institute of Technology; WILLIAM H. MECKLING, University of Rochester; WILLIAM K. JONES, Columbia University

8:00 P.M.

Richard T. Ely Lecture

Chairman: ABRAM BERGSON, Harvard University

Speaker: NICHOLAS GEORGESCU-ROEGEN, Vanderbilt University

Monday, December 29, 1969

8:30 A.M.

Basic Data for Policy Decisions: Public and Private

Chairman: THEODORE YNTEMA, Oakland University

Papers: CHARLES L. SCHULTZ, Brookings Institution; DONALD TURNER, Harvard University; LEONARD SILK, Brookings Institution

History of Economic Thought

Chairman: CRAUFORD GOODWIN, editor, *History of Political Economy*

Papers: MARTIN BRONFENBRENNER, Carnegie-Mellon University; NORFLEET W. RIVES, Duke University

Discussants: WILLIAM R. ALLEN, University of California, Los Angeles; VINCENT J. TARASCIO, University of North Carolina

The International Firm and Efficient Economic Allocation

Chairman: CHARLES P. KINDLEBERGER, Massachusetts Institute of Technology

Papers: ROBERT E. BALDWIN, University of Wisconsin; JACK BARANSON, International Bank for Reconstruction and Development; STEPHEN HYMER, Yale University

Discussants: THOMAS WILLETT, Harvard University; STEFAN H. ROBOCK, Columbia University; RONALD MCKINNON, Stanford University

Basic Data for Policy and Public Decisions: Technical Aspects

Chairman: JOHN R. MEYER, National Bureau of Economic Research

Papers: GUY H. ORCUTT, Urban Institute; THOMAS JUSTER, National Bureau of Economic Research; CHARLES C. HOLT, University of Wisconsin

Discussants: WILLIAM F. SHARPE, University of California, Irvine; RICHARD RUGGLES, Yale University; EDWIN KUH, Massachusetts Institute of Technology

10:30 A.M.

Economics of Education (Joint session with the Industrial Relations Research Association)

Chairman: CLARK KERR, Carnegie Commission on the Future of Higher Education

Papers: ROY RADNER, Churchill College, England, and LEONARD S. MILLER, State University of New York at Stony Brook; W. LEE HANSEN, University of Wisconsin; JOSEPH KERSHAW, Ford Foundation, and ALEX MOOD, University of California, Irvine

Discussants: WILLIAM G. BOWEN, Princeton University; ALLAN M. CARTER, New York University; KARL A. FOX, Iowa State University; W. ALLAN WALLIS, University of Rochester

Current Economic Policies: Their Appropriateness and Effectiveness (Joint session with the American Agricultural Economics Association)

Moderator: DALE E. HATHAWAY, Michigan State University

Panel: WALTER W. HELLER, University of Minnesota; D. GALE JOHNSON, University of Chicago; JOHN A. SCHNITTKER, Kansas State University; HENRY C. WALLICH, Yale University

12:30 P.M.

Joint Luncheon with the American Finance Association

Chairman: LAWRENCE S. RITTER, New York University

Speaker: ROBERT V. ROOSA, Brown Brothers Harriman & Co.

2:30 P.M.

The Economy of Cities

Chairman: WASSILY LEONTIEF, Harvard University

Papers: JANE JACOBS,³ Consultant, Toronto; BARBARA BERGMANN,⁴ University of Maryland; ANTHONY DOWNS,³ Real Estate Research Corporation

Discussants: KARL A. FOX, Iowa State University; DAVID S. MUNDEL, Massachusetts Institute of Technology; SAM BASS WARNER, JR., University of Michigan; ROBERT C. WEAVER,¹ Bernard M. Baruch College

Invited Doctoral Dissertations

Chairman: RICHARD RUGGLES, Yale University

Papers: E. ROY WEINTRAUB, Rutgers University; PEDRO BELLI, University of California, Riverside; MYRON G. MYERS, Rutgers University; ROBERT J. ANDERSON, Purdue University

Discussants: LIONEL W. MCKENZIE, University of Rochester; DAVID KENDRICK,¹ Stanford University; LOWELL S. HARDIN, Ford Foundation; RONALD JONES, University of Rochester

Investment Behavior

Chairman: DALE W. JORGENSEN, Harvard University

Papers: CHARLES W. BISCHOFF, Yale University; LAURITS R. CHRISTENSEN, University of Wisconsin; ROBERT E. LUCAS, Carnegie-Mellon University

Discussants: ARNOLD C. HARBERGER, University of Chicago; VERNON SMITH, University of Massachusetts; RALPH B. BRISTOL, JR., Treasury Department

John R. Commons (Joint session with Omicron Delta Epsilon)

Chairman: ERVIN K. ZINGLER, University of Houston

Papers: GERHARD TINTNER, University of Southern California; ALAN BROWN, University of Indiana

Discussants: E. BRYANT PHILLIPS, University of Southern California; PETER S. ELEK, Villanova University

8:00 P.M.

Presidential Address

Chairman: GOTTFRIED HABERLER, Harvard University

Presidential Address: WILLIAM J. FELLNER,⁴ Yale University

9:15 P.M.

Business Meeting

Tuesday, December 30, 1969

8:30 A.M.

The Teaching of Economics

Chairman: JOHN M. KUHLMAN, University of Missouri

Papers: RICHARD EDWARDS, ARTHUR MACEWAN, and OTHERS, Harvard University; W. LEE HANSEN, ALLEN C. KELLEY, and BURTON A. WEISBROD, University of Wisconsin; BEN LEWIS, Ford Foundation

Discussants: PAUL SWEETZ, Editor, *Monthly Review*; PHILLIP SAUNDERS, Carnegie-Mellon University; JOHN R. COLEMAN, Haverford College

³ Printed in the September, 1969, *A.E.R.*

⁴ Printed in the March, 1970, *A.E.R.*

AMERICAN ECONOMIC ASSOCIATION

Distribution Issues: Trends and Policies**Chairman:** RICHARD A. MUSGRAVE, Harvard University**Papers:** EDWARD C. BUDD, Pennsylvania State University; LESTER C. THUROW, Massachusetts Institute of Technology; ROBERT J. LAMPMAN, University of Wisconsin**Discussants:** MICHAEL K. TAUSSIG, Rutgers University; T. PAUL SCHULTZ, RAND Corporation; STEPHAN MICHELSON, Harvard University***Forecasting and Policy Evaluation Using Large-Scale Econometric Models*** (Joint session with the Econometric Society)**Chairman:** JON CUNNINGHAM, Ohio State University; LAWRENCE R. KLEIN, University of Pennsylvania**Discussants:** KARL BRUNNER, Ohio State University; STEPHEN M. GOLDFELD, Princeton University; ROBERT A. GORDON, University of Chicago; DANIEL SUITS, University of Michigan***Graduate Student Papers*** (Sponsored by the *American Economist*)**Chairman:** JOHN D. GULLFOIL, New York University**Papers:** BENNETT HARRISON, University of Pennsylvania; ALAN DEARDORFF, Cornell University; DONALD CAMPBELL, Princeton University; THEODORE KEKLER, Massachusetts Institute of Technology**Discussants:** ELIZABETH DURBIN, Columbia University; ALFRED BROADDUS, Indiana University; BRUCE D. RESEHN, City University of New York; THOMAS CONOSCENTI, New York University

10:30 A.M.

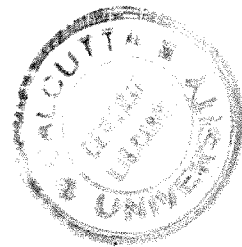
Economics of Imperialism**Chairman:** PAUL M. SWEEZY, Editor, *Monthly Review***Papers:** RICHARD WOLFF, City College of New York; THEOTONIO DOS SANTOS, University of Chile; HARRY MAGDOFF, New School for Social Research**Discussants:** STEPHEN HYMER, Yale University; VICTOR PERLO; ARTHUR MACLEWAN, Harvard University***New Dimensions in Natural Resources*****Chairman:** HAROLD J. BARNETT, Washington University**Papers:** FRANCIS T. CHRISTY, Resources for the Future; JEROME ROTHENBERG, Massachusetts Institute of Technology; RAYMOND VERNON, Harvard University**Discussants:** RICHARD R. NELSON, Yale University; CHANDLER MORSE, Cornell University; MARSHALL HALL, Washington University***Dynamics of Income Distribution: Poverty and Progress*****Chairman:** HAROLD W. WATTS, University of Wisconsin**Papers:** CHARLES E. METCALF,¹ University of Wisconsin; JAMES N. MORGAN, University of Michigan; and JAMES SMITH, Pennsylvania State University**Discussants:** ARTHUR M. OKUN, Brookings Institution; WILLIAM BRANSON, Princeton University; LAWRENCE R. KLEIN,¹ University of Pennsylvania; MICHAEL J. PIORE, Massachusetts Institute of Technology

2:30 P.M.

New Ideas in Pure Theory**Chairman:** WASSILY LEONTIEF, Harvard University**Papers:** HERBERT SCARF,¹ Yale University; LLOYD SHAPLEY,² RAND Corporation; ROY RADNER, Churchill College, England**Discussants:** RUTH MACK, Institute of Public Administration; KENNETH J. ARROW, Harvard University; TROUT RADER, Washington University***Innovation and Public Enterprise: East and West*** (Joint session with the Comparative Economics Association)**Chairman:** ABRAM BERGSON, Harvard University**Papers:** JOSEPH BERLINER, Brandeis University; WILLIAM G. SHEPARD, University of Michigan; JOHN SHEAHAN, Williams College**Discussants:** DAVID GRANICK, University of Wisconsin; ABRAHAM P. LERNER, University of California, Berkeley; CHARLES P. KINDLEBERGER, Massachusetts Institute of Technology***Manpower Dimensions of Economic Growth*** (Joint session with the Industrial Relations Research Association)**Chairman:** GEORGE O. SELTZER, University of Minnesota**Papers:** JOHN C. SHEARER, Oklahoma State University; EVERETT M. KASSALOW, University of Wisconsin**Discussants:** HARRY G. JOHNSON, University of Chicago; NOAH M. MELTZ, University of Toronto; HERBERT S. PARNES, Ohio State University

THE purpose of the American Economic Association, according to its charter, is the encouragement of economic research, the issue of publications on economic subjects, and the encouragement of perfect freedom of economic discussion. The Association as such takes no partisan attitude, nor does it commit its members to any position on practical economic questions. It is the organ of no party, sect, or institution. Persons of all shades of economic opinion are found among its members, and widely different issues are given a hearing in its annual meetings and through its publications. The Association, therefore, assumes no responsibility for the opinions expressed by those who participate in its meetings. Moreover, the papers presented are the personal opinions of the authors and do not commit the organizations or institutions with which they are associated.

HAROLD F. WILLIAMSON
Secretary



Number 71 of a series of photographs of past presidents of the Association



William Fellner

Trends in the Activities Generating Technological Progress

By WILLIAM FELLNER*

Much of mankind's accumulated knowledge performs functions other than that of increasing our command over goods and services in the usual sense of these words. The value of such knowledge is not instrumental in an economic-technological context. It is equally clear that contributors to the growth of knowledge have all along been motivated largely by a desire to improve their understanding of the world into which they were born. "From science to engineering and from engineering to more goods and services" would be an unduly drab account of man's quest for knowledge.

Yet in societies of various types, financial and other incentives have become established for directing inventive and educational abilities to specific tasks carrying economic promise. Moreover, for several centuries the Western conception of knowledge that is valuable for its own sake has been appreciably influenced by the ideals of natural scientists and mathematicians. It is a fact of crucial importance that this has created a cultural environment in which a reasonably high degree of correlation may be found between the accumulation of knowledge per se and subsequent technological applications. The methods of measuring the economic yield of new knowledge and of measuring the

trend of that yield have many inevitable limitations reflecting the indirectness of the relation of the knowledge-acquisition process to its economic consequences. My paper will be concerned with these yields and trends.

The inputs which in this study will be regarded as progress-generating will be so defined because they increase the *economic* productivity of inputs at large. Our progress-inputs could be viewed as producing intangible capital that is instrumental to the production of goods and services and also intangible *consumer* capital serving as a source of direct satisfaction. But this is merely a *simile* because the properties of this intangible investment militate against treating it as output in the sense proper. The "intangible investment" results in an increase of the value of terms which in static versions of neoclassical production functions would be *parameters*. The progress-generating inputs of each period may then be viewed as producing additional output indirectly, via their effect on such terms of otherwise conventionally defined production functions. In my appraisal, interpreting the progress-generating inputs of any period, and their immediate results, in this distinctive fashion describes the most convenient way of separating them out for specific analysis.

As I shall show later, however, very similar results are obtained in an analytical framework of different character in which *practically all* long-run increase in output per man-hour is interpreted as developing from additional per capita knowl-

* Presidential address delivered at the eighty-second meeting of the American Economic Association, New York, New York, December 29, 1969.

This paper has developed from research in which I have benefited from the valuable assistance of Mr. John Zerbinis.

edge.¹ Even what in the usual neoclassical model is considered mere factor-substitution is in this alternative framework viewed as implying the *invention* of new types of goods, viz. of less labor-intensive equipment. I will explain later why I will merely keep an eye on the conclusions obtainable from this construction which is neither neoclassical nor Cambridgeian, and why the main part of my analysis will move in the framework usually referred to as neoclassical.² Objections raised

¹ In the case of the United States "practically all," because here terms-of-trade effects, the consequences of improved diets, etc. play no major role. There remains an ambiguity in connection with the consequences of intersector shifts, on which see the discussion of the two Kendrick measures in Section III.

² To illustrate with the simplest kind of Cobb-Douglas function applicable to the problem, as well as this is possible for a single period ($t=n$):

$$Q(n) = F \left(\sum_0^n L_R, \sum_0^n K_R, L_B(n), K_B(n), L_A(n), K_A(n) \right) L_C^{\alpha}(n) K_C^{\beta}(n)$$

L_R, K_R = research-type inputs, in the broad sense.

$L_B(n), K_B(n)$ = education (knowledge-distribution) inputs which were applied in the past, or are being applied in period n , to the labor force of period n , expressed *per capita* of this labor force.

$L_A(n), K_A(n)$ = inputs required for changing over to improved industrial methods (*mainly* the obsolescence-component of capital-replacement).

$L_C(n), K_C(n)$ = those inputs of period n not accounted for in the argument of F , except that the education inputs applied in period n to persons who are not yet members of the labor force belong neither in $L_C(n), K_C(n)$ nor of course in the argument of F .

The algebraic form and the parameters of F depend on factors not formalized in the equation, such as abilities and habits of the population; policies other than those expressing themselves in the R and E inputs; various inputs in foreign economies; the relative weight of that part of the old knowledge which remains useful when new knowledge is acquired, etc. It will be seen later that the convenient definition of the rate of technological progress here is $d \ln F/dt$.

In this model avoidance of the implication of economies of scale would require $\beta < 1 - \alpha$, instead of the conventional $\beta = 1 - \alpha$, but it may be preferable to face the problem of such economies as one giving rise to the qualifications discussed in the opening paragraph of Section III.

The progress-generating inputs should be defined so that if the variables in the argument of F assume the

against the neoclassical models are quite inconclusive in their present form, because the question is not whether the assumptions underlying such models are "realistic" (a good photograph of reality), but whether in the real world these assumptions are violated in such respects and to such an extent as to render the analytical results misleading. This the critics have so far not even *tried* to show.

I. A Ratio of Benefit-Stream to Cost and the Corresponding Average Social Rate of Return

The costs of generating progress consist of the costs of *producing new knowledge*, and the costs of *increased per-capita knowledge-distribution*, and of the extra-costs of *changeovers* from the use of old to the use of new knowledge. This seems a reasonable way of looking at our problem, because if knowledge in use remained constant per capita, the rate of progress would be

values to which zero progress-generating input corresponds, then $d \ln F/dt = 0$ (in which case, however, L_C and K_C would be larger than at positive values of the progress-inputs and of this derivative). Hence the progress-generating inputs for *period n* consist of the period's additions to the two \sum terms; and of the increase of L_B and K_B expressed per capita of the labor force of period n , as compared to the immediately preceding period's L_B and K_B expressed per capita of that period's labor force; and of $L_A(n)$ and $K_A(n)$. In principle, one of the magnitudes representing a progress-input (increased per capita education) could turn out to be negative, and avoidance of this awkwardness would then call for some adjustments of terminology, but in our context this may be disregarded.

If the methods of measuring progress call for limiting Q to the private sector (usually including government-owned *enterprise*), then L_C and K_C should also be so limited, but the argument of F should nevertheless include also the public-sector inputs of the type specified in the lines under the equation.

The larger cost-base to be defined in Section II involves "charging" to the progress generated in a period all progress-generating inputs for the period; the smaller cost-base to be defined in Section II involves charging merely the "profit-oriented" progress-inputs, which in turn implies that the costs of the other progress-inputs are recovered outside the market process and that the progress-account should therefore be "credited" with these nonmarket values.

zero. Whether we wish to distinguish the specific costs of changeovers as a separate subcategory is a terminological question, since technological knowledge is not even truly *acquired* before it is introduced on an industrial scale. But at any rate the changeover-costs belong among the costs of progress, and they are quantitatively significant though they are frequently overlooked when benefits are weighed against costs. The costs of progress so conceived are the costs of those inputs which in this paper are referred to as "progress-generating." These inputs become combined with *human talent* which is no easier to define and is certainly no less essential than are the Ricardian "original" and "indestructible" powers of the soil.

We may express, as a *proportion* of the value of a year's output, the cost of the goods and services representing the progress-generating input for the year; and on the benefit-side, we may set against this cost the *proportion* in which technological progress increases the output flow annually.³ On specific simplifying assumptions, the progress-induced proportionate addition to output will remain permanent and independent of the output-base. With a deduction on the productivity-side to allow for the continuing need to keep per capita education at any once-attained higher level,⁴ we then arrive at the ratio of an economic benefit-stream to cost; and after multiplying by 100, we obtain a magnitude that comes close to what the economist may consider a measure of the average social rate of return from the progress-generating activities of the period in question.⁵ Such a concept of an *average*

rate of return can not, of course, guide us to the core of the ultimately relevant problems of resource-allocation, but the concept will nevertheless be found useful. Even on assumptions assuring constant proportionate additions to the output-flow the delayed start of the benefit stream will subsequently require allowances in the nature of discounts, yet for reasons explained in Appendix A and in Section VI, this difficulty will not prove critical.⁶ We will also have to try to get rid of the effect of cyclical and erratic forces on our yearly data.

The simplifying assumptions required for defining average social rates of return in terms of *constant* proportionate additions to output are indeed specific. But analysis based on these assumptions has led to reasonably realistic results in a good many other respects. I shall make these assumptions which should assure at least rough comparability of the benefit-cost ratios so obtained with analogous ratios for physical capital formation on a more-or-less given level of technological knowledge.⁷ The "stepping up" of the *absolute* contribution of technological progress when future inputs (other than progress-generating ones) raise the output-base must not be allowed to destroy the comparability of progress-returns with the returns from investment in the usual sense. This must be watched in formulating the assumptions needed for the justification of such comparisons.

also be expressed by the statement that the ratio of the discounted benefit-stream to costs is greater than one.

³ One of the main reasons for not becoming very disturbed by this difficulty is that the rates of return from progress-generating activities will be compared with the rate from physical investment "at a given level of technology," and in these comparisons allowance needs to be made *also* for another factor which can be interpreted as the equivalent of a lag working *against* physical capital formation (i.e., in favor of the progress-activities). For another reason see Section VI.

⁷ The reason why I say "more or less" is explained in the first paragraph of Section III.

³ The definition of output should be the same on the two sides of the ledger.

⁴ It will be seen, however that this deduction is called for only to the extent that education is charged to the progress-account, i.e., is not netted out (Cf. Sections II and IV).

⁵ If we find that the resulting rate of return is higher than the "normal" rate in the economy, then this may

The required simplifying assumptions are reasonably well satisfied in the macro-economic Cobb-Douglas framework. The CES model also acquires the needed properties if we work into it a mechanism of induced invention and of distributive-shares equilibrium.⁸ This is explained in Appendix A. In Sections III and IV we shall return to the question *how much* similarity of the real world with such models is implied in my paper.

I should add that on both sides of my benefit-cost ledger I will have to disregard problems of considerable importance which either cannot be articulated sufficiently well, or, if capable of articulation, can be appraised only in very vague terms. I will not be able to consider the question whether all-too-rapid change does not reduce our well-being "in a fundamental sense," or, on the other hand, whether at a stagnant or near-stagnant level of technology the institutions of Western nations would be workable at all. Nor will I try to explore the relationship of technological progress to population growth and to its specific consequences.

II. Two Versions of the Social Benefit-Cost Ratio, and their Meaning

A substantial part of the progress-generating costs of each period is incurred by decision makers subjected to the criterion of private economic yield. The progress-costs falling in this category consist mainly of the costs of privately financed research and development (*R & D*) and of the extra costs of changing over to new industrial processes. Though we are concerned here with the problem of social yield, we should obviously include these private yield-oriented costs in the cost-base of our benefit-cost ratio.

⁸ See my December 1967 article which is listed in the Reference section, as are the particulars of the other references. The bibliography attached to that article lists earlier writings of several authors on this subject (including some by the present writer).

Another large part of the progress-generating expenses is incurred by agencies not institutionally subjected to criteria of economic yield. The government-financed *R & D*⁹ and that financed by other non-profit institutions belong here, and so does almost all of the cost of increased per capita education.¹⁰ This does not mean that the nonprofit institutions incurring these costs are under no political or quasi-political¹¹ pressure to keep costs down per unit of service; what it does mean is that these costs create net income by definition, i.e., that their net-income-creating property does not depend on whether they generate offsetting revenues. Nonprofit institutions, most prominently illustrated by governments, but illustrated also by private teaching organizations, can have deficits but these deficits are not usually viewed as negative income such as would cancel any part of the incomes created by the institutions. In the language of national accounting, the services of these institutions are *valued at cost*, though, when appraising the worth of the services, any individual may find this valuation unconvincing. In this regard such costs are treated differently from those of firms, e.g., differently from that *R & D* component which is financed by industry.¹² The

⁹ In the United States the greater part of the government-financed *R & D* is performed on contract by industry. The remainder of the government-financed component is performed by the government itself and by other nonprofit institutions (see Appendix D).

¹⁰ Training on the job in the business sector, and sales of books, periodicals and newspapers to consumers represent exceptions. These are institutionally profit-oriented expenditures. They do not weigh heavily in the sum total of education costs in the broad sense (and, of course, these items belong here *only* if education costs are interpreted in the broad sense, as they should be for our purpose).

¹¹ By quasi-political pressure, I mean the pressure to which trustees and administrators of private nonprofit organizations are exposed and which they in turn exert on the staffs of these organizations.

¹² Industry finances also some of the research performed by universities and other nonprofit organizations but more than 98 percent of the *R & D* financed by industry is also performed by it. (Continued)

costs of industry show in losses—in deductions from net income—unless they are recovered through revenue from sales in markets.

This difference in accounting methods corresponds to differences in the strength of the argument for including the two types of progress-generating expense in the costs that make up the denominator of our ratio of benefit-stream to costs. Where the expenses need not be recovered through revenues generated by them, groups of decision makers were given legal authority to proceed on the assumption that the knowledge so produced or distributed has considerable *nonmarket value*, i.e., value which is not measured as part of the economic benefit-stream. As a rule this is true *even* where a nonprofit organization¹³ makes its services available through sales in markets, as is the case in a subsector of the education sector. Even these nonprofit organizations are heavily subsidized *and/or* the acquisition of the *type* of service they offer has been made compulsory by a political decision limiting the *market* decision of the customers to the choice of a source of supply. The question arises therefore whether the practice of valuation at cost to the nonprofit institution, without regard to recovery of the cost through revenue, should for our purpose be taken at its face value in the sense that we *net out*, and thus *exclude* from the cost-side of our social benefit-cost ledger, the costs of all items that the national accountant treats as if they created equivalent values by defini-

tion. In a preliminary step let us do so, but before trying to appraise the merits and deficiencies of this version of our ratio, let us take a look also at another way of going about the matter.

We may, alternatively, decide to *include* in our cost-base the costs of nonprofit institutions along with those of the profit-seeking ones and thus obtain a ratio based on *all-inclusive* costs. This ratio will, of course, come out lower than the first.

We must remember, however, that even the cost-base which we call "all-inclusive" excludes items which are interpreted here not as costs of progress-generating inputs but as expenditures *influencing the effectiveness of such inputs*.¹⁴ In the first place, we shall have to limit ourselves to the analysis of American data; hence the costs incurred abroad belong among those automatically excluded. Secondly, medical expenses, other health-related components of the consumer budget, and types of public investment bearing on the effectiveness of progress-generating inputs are excluded (netted out) *even* from our larger cost-base because it is implied that these expenses become justified by the specific output which they buy, *aside from* their effect on productivity-gains. The same is not "automatically" taken for granted for any part of the input we define as progress-generating. No part of that input is excluded from our larger cost-base, and while part of it *is* excluded from our smaller base, we emphasize that both bases require attention in our analysis. Thereby we leave a range open for personal (subjective) valuation of the contribution of progress-generating inputs *aside from* their contribution to measurable productivity gains.

The proposition suggests itself that the

In *R & D* statistics the concept of "industry" is intended to include the whole business sector. In principle individuals should be included even if they have not founded "firms" but while the role of some such individuals has been of considerable importance in the history of technological progress, their weight on the cost-side of our ledger is very small. See John Jewkes, David Sawers, and Richard Stillerman. The costs of some other strategically significant inputs are also small (see Appendix C).

¹³ Government-owned *enterprises* are not included in this concept.

¹⁴ In our framework this distinction coincides with the distinction between the argument of a function and factors influencing the algebraic form or the parameters of the function. See fn. 2.

average social benefit-cost ratio which a person is willing to accept as relevant to his behavior as a voter, adviser, or administrator is unlikely to fall significantly outside the range bounded by the two ratios we have defined. Our lower ratio involving our all-inclusive cost-base completely disregards the value of new knowledge per se and also its instrumental value for the performance of services made available to the public through channels other than those of the market. Our higher ratio involving merely the costs of the institutionally profit-oriented progress-inputs implies that the given scale of the progress-generating activities of governments and other nonprofit organizations would be justifiable even if these activities had no favorable effect whatever on the market sector; in view of the actual scale of these activities in the contemporary American economy, such a judgment would, I think, be interpreted by most individuals as describing valuation *at, or near*, the upper limit of the range of reasonable valuations. Thus, for what we may consider typical valuation, our lower boundary seems convincing *as such*, and our upper boundary seems not far off the mark *as such*.

Can these propositions concerning alternatively computed ratios of social benefit stream to cost provide guidance for the appraisal of the results of the American progress-generating activities? I suggest that the answer is in the affirmative, though, it must be recognized that an affirmative answer implies a rather modest interpretation of what it means to be guided in one's behavior by an estimated rate of return. Such a limitation is inherent in any analysis that focuses on *average* rates of return for successive periods. Yet there is good reason for choosing this focus in the central part of the present analysis.

We are concerned here with activities for which the *marginal* social returns, and hence the ultimately relevant social allo-

cation criteria, are not capable of being estimated by examining private *or* national accounts. Even with regard to the progress-generating inputs financed by profit-seeking institutions, significant external effects and shifts in competitive positions create differentials between social returns and the private marginal returns of the innovating firms. For the diffusion of knowledge in the market sector reduces their profits, while gains in competitive position achieved during the intervening period may prove enduring. These gains increase the private returns. At the same time, the spillover effect of the progress-activities originating in the nonprofit institutions expresses itself in part in private market-returns. As for the progress-generating activities of nonprofit institutions, obviously there exists no way of approximating from accounting data the value of their nonmarket contribution *or* that of their spillover effect on the market economy. When appraising marginal social returns each individual is entitled to his own standards, and where the problem is one of politically determined inputs, individual valuations will not lead to the quantity-adjustments which would work toward equalization of returns at the margin.¹⁵

In such circumstances only one statement can be made about the *marginal* criteria of desirable allocation on a reasonably "objective" (or general) level of discourse. The statement is that *once a very large number of individual preferences concerning the ranking of the specific objectives of the progress-activities have been bridged*

¹⁵ As was pointed out above, in some sectors of the progress-activities—particularly in a subsector of "education"—political-type decisions, which greatly influence the aggregate quantity of resources used for a general purpose, nevertheless leave room for the individual's market-choices as to the specific quantities of service acquired from alternative suppliers. I consider extension of the range of such choices desirable but whatever the range is (or might become) its existence does not contradict the statements in the text.

by *political or quasipolitical compromise*, it is in the common interest of the community to avoid waste. The content of the political compromise plays a significant role in shaping the waste-avoidance problem with which the private as well as the public sector is faced. Given some specific compromise on the priority-rankings, the avoidance of waste requires the solution of technical problems that are far from trivial from a professional point of view. Yet any *substantive* suggestion on how to satisfy the marginal criteria of socially desirable allocation for the progress-inputs inevitably implies distinctly personal preferences which need to be reconciled with the preferences of others. Preferences of this sort are essentially preferences of an individual for one variety of political compromise ("as if" utility function) rather than another.¹⁶

This very high degree of subjectivity does not attach to all substantive propositions that may be developed on the yield of the progress-effort. It is possible to suggest substantive propositions not only about the average social rate of return from each period's progress-activities but also about the algebraic sign of the difference between the average and the marginal social rate to which a much lesser degree of subjectivity attaches. *Even* if an individual were in a position to exert decisive influence on the composition of the progress-generating inputs—or if for any reason he identified himself fully with politically determined priority-rankings—he would still be apt to find a considerable gap between the average and the marginal social rates from the progress-activities of a period, provided that these activities are pursued on a significant scale. The gap between average and marginal rates develops because high talent shades over into lesser talent, and because the solution

of problems is an inherently sequential process. This is an essential property—we may perhaps say a *Schumpeterian* property—of the progress-generating activities. By way of "transitivity," the following propositions follow from the average-marginal gap in the progress activities:

a) Each person has reason to consider it a *necessary* condition of his approval of a set of allocative decisions that by his standards the *average* social rate of return from the progress-activities should exceed the *marginal* social rate of return from physical investment "at a given level of technological knowledge";

b) the necessary condition requires also that the average rate of progress-return should exceed the marginal market rate from the current flow of physical investment, since we may assume that this market rate is no higher than the marginal social rate from physical investment;

c) finally, no undue stretching is involved in extending this condition to the requirement that the average rate of progress-return should exceed the rate of return on the capital stock—say, the algebraic product of the share of capital in income with the ratio of yearly output to capital¹⁷—since we may assume that normally this rate is not very different from the marginal market rate on physical investment.

It therefore seems reasonable to conclude that from each individual's point of view it is a necessary condition of good allocation that by his standards the average social rate of return from the progress-generating activities of the successive periods should tend to exceed the profit rate on physical capital, i.e., should exceed the rate which in traditional theory is interpreted as the net marginal productivity

¹⁶ Subjective probability judgements also need to be reconciled.

¹⁷ On statistics relating to corporate profits, see fn. 47. The algebraic product described in the text above gives an *approximation* to a magnitude which is useful for our purpose. See Section I and Appendix A.

of capital. About the relationship between the average rate from the progress-activities and the conventionally estimated marginal product of capital it is possible to make reasonably "objective" numerical statements, though unfortunately not very precise ones, because some statements that can be made about the *average* rates from the progress-activities imply merely value judgments that are very widely shared.

In particular, a great many individuals would agree that an unduly strict criterion of overall social profitability, and hence of required average social rates of return, is set by charging the full cost of the knowledge-acquisition activities of governments and of teaching and research institutions to measured productivity trends, *as if* the nonmarket contribution of these activities were zero. At the same time, most individuals would, I think, agree that it is useful to inquire into the question of average social rates of return also by valuing the nonmarket contribution of the politically or quasi-politically determined progress-inputs at full cost to the nonprofit institutions *as if* no favorable effect on the market-sector were required to justify these inputs on their present scale. I surmise that the valuation convincing to most persons would be located either within, or not significantly outside, the range described by these two frames of reference for valuation. In this sense, our two benefit-cost ratios describe a range which should be of some use to most decision makers, particularly because the usefulness of the framework is admittedly limited. A verdict of social profitability as expressed by adequate average rates of return is compatible with anyone's personal judgment that the progress-activities could be made even more profitable by changing the controversial content of political compromises concerning the composition of the inputs. While the present paper will not lose sight of these highly

controversial questions involving marginality, I will try to separate these from the problems that can be analyzed on a reasonably "objective" level. In practice, this objectivity becomes reduced by the need to fill in gaps in the available data by operations that at times become very speculative.

More ambitious contributions of other authors are revealing in many ways but they do not remove the limitations we need to observe in our present context. For example, from empirical investigations one may conclude that additional education—education *at the margin*—is profitable by market criteria alone, even without regard to socially desirable external effects and to the nonmarket value which is acquired.¹⁸ This is a significant result, and it has important applications, but it cannot be put to use in our present analysis. Additional education is profitable very largely because our technology is becoming increasingly complex, a process which in turn presupposes additional education. Recognition of this takes us back to the problem of the social worth of the activities that lead to the increasing complexity of our technological and organizational methods.

III. *The Residual: A Roundabout Way of Measuring Productivity-Increases*

The method here to be used for estimating productivity gains leads to higher estimates than do some procedures that have occasionally been suggested. Yet I feel that the method needs to be defended not so much against methods that attribute a smaller component of growth to progress, but against methods that attribute to progress an even larger growth-component. As concerns the United States, an argument could be made for the view that *practically all* long-run increase in output per man-hour (or possibly more than this) involves invention and/or

¹⁸ See particularly Gary S. Becker.

additional knowledge-distribution because even insofar as the neoclassical approach attributes a growth-component to mere factor-substitution—i.e., insofar as physical investment exceeds the equivalent of mere “widening”—such investment presupposes the invention of new equipment of lesser labor-intensity.¹⁹ *Practically all* long-run increase of output per man-hour (and even the prevention of some decrease resulting from land-scarcity) requires at least some amount of additional per capita knowledge which may or may not express itself in the production of new types of equipment²⁰; the reverse proposition that making use of new ideas requires new capital goods is hazy because new capital goods result from new ideas applied to already existing types of equipment. If one wanted to follow through consistently the implications of the conception that all long-run increase in man-hour output reflects technological progress, one would arrive *at even higher* productivity-gains than those which we shall assume (though at more costly ones); the progress-generated productivity-gains so defined would show a rising tendency, as will ours; and these progress-gains too would point to high but somewhat declining *rates* of social return, similar to ours.²¹ We will not follow through all

the implications of this “inclusive” conception of progress though *one* of its several implications will be carried over into the framework I am using because in mine, too, economies of scale will become merged with “progress.” My main reason for not following through all the implications of the most inclusive conception is that some possible ways of increasing man-hour output require new ideas only in a very modest sense of this term, i.e., require merely somewhat trivial new ideas (resulting in slow changes of the physical composition of the capital stock). This is what in my appraisal justifies the “neo-classical” procedure of attributing part of the increase in output per man-hour to changes in input-proportions *aside from* technological progress. But this way of looking at the problem still leaves one with a preference for principles that lead to higher progress-estimates than do the principles suggested in some other contributions to growth theory.

My analysis will rely on the method of the Residual as it is employed by John Kendrick for measuring the proportionate rate of increase of productivity. From this productivity-increase I will derive the magnitude I consider the benefit resulting from the progress-generating costs.

As is well known, the method of the Residual attributes to technological advance *that* increase in output which is *not* attributed to the increase in physical inputs given their *base-period productivities*. The measurement of productivity assumes measurement of the price-corrected market

¹⁹ This does not, however, imply that progress-generating inputs (in our sense) and investment in less labor-intensive equipment come in fixed proportions.

²⁰ For qualifications see fn. 1. In the modern American context these will be disregarded.

²¹ On such an “inclusive” interpretation of progress one would have to charge our progress-generating inputs, and also the net capital formation *other than the equivalent of mere “widening,”* to the increase of man-hour output (presumably with a crude allowance for the fact that the consequences of the relative scarcity of land also become offset along the growth path). The average rates of return on this bundle of physical investment and “progress-generating inputs” (in our sense) would be *intermediate* between what in the main part of this paper will be regarded as returns from the progress-inputs and what will be regarded as returns from physical investment “at a given level of technology.” For example, the lowest of the calculated rates of progress-return in fn. 40,—clearly an underestimate

of the true rate—is 13 percent; the corresponding “lowest rate” (underestimate) in the framework implying the “inclusive” conception of progress may be said to come out at slightly less than 12 percent. Given this same “inclusive” conception of progress, the rate on that part of the physical investment which exceeds the equivalent of mere widening would be *zero* in the absence of all progress; on the other hand, in the absence of net capital formation the rate of return from progress-generating inputs would *not* be zero (see Section IV, 2 below).

value of output, and in the United States this is practically the same thing as to say that we are concerned here with a Residual observed as a proportion of the *privately produced* output.²² The procedure has the implication that the analysis of the determinants of technological progress is as yet in a tentative stage, so that if for example *R & D*, the increase of per capita education, the extra-inputs needed for industrial changeovers, and some other items to be discussed in this paper make up the identifiable progress-inputs (as I suggest they do), we cannot at present expect good results from regressing output on all these specific inputs along with the conventional inputs specified in the production functions. I feel convinced that realism does require recognizing this limitation but that this does not mean that we are left with no good argument for relating the Residual to the inputs and costs I defined as progress-generating.

Regression analysis of cross-sections of individual industry groups and of firms enabled investigators to obtain confirmation of the hypothesis that productivity-increases, as estimated by the method of the Residual, are positively correlated with *R & D*-intensity.²³ On the other hand,

²² However, the base in relation to which the productivity-increase is measured is somewhat inflated as a result of the fact that the private sector includes some activities for which no productivity-increase can be measured (such as the activities of private nonprofit organizations, and paid household work); and it includes some activities for which the output indices actually used do not measure productivity-increases adequately. A minor terminological awkwardness to be noted here is that government-enterprise, in contrast to general government, belongs in the "private" sector. See also fn. 33.

²³ Kendrick (pp. 182-186) has found a statistically significant correlation between the increase in his total factor productivity and the "*R & D*-intensity" of specific industry groups. His discussion of this problem is based partly on his own research and partly on that of Nestor Terlecky.

Jora Minasian in his May 1969 article, and in his earlier work there quoted, reports on a positive relationship between *R & D* and productivity-increase in

spillover effects from the public into the private sector, the diffusion of knowledge within the private sector, and the narrowness of the concept of *R & D*-intensity as compared to our concept of progress-generating input make it desirable to ask the question of the relationship of the Residual to progress-inputs also in broader *intertemporal* terms. When the question is so posed, it calls for exploring whether the time-sequence of productivity increases, as measured by the Residual, bears an understandable relation to the time-sequence of progress-generating inputs. We shall see that this is indeed the case, but in such work there is at present only limited room for regression analysis and for "testing" in the technical sense. This is because for some purposes for which time-sequence comparisons of productivity increase with inputs are used it is advisable to rely on a rather small number of typical values of the variables for successive periods of considerable duration; high serial correlation and lags reduce the usefulness of data reflecting *yearly* changes.²⁴ Emphasis will

a cross-section of chemical firms. He regressed Value Added on labor, capital and *R & D*. His estimate of the rates of return from *R & D* disregards a substantial proportion of the progress-costs with which we shall be concerned in this paper, but what matters here is the existence of a positive relationship between *R & D* and productivity-increase. However, the fact that on the cost-side merely *R & D* enters into Minasian's analysis when he estimates rates of return may provide part of the explanation why these rates are hard to reconcile with the results of a survey (questionnaire study) on which Edwin Mansfield reported also in the May 1969 *Papers and Proceedings*. Firms must be acutely aware of the changeover costs and of various other costs to be considered in our analysis.

²⁴ This, however, is not all "black and white." For example, regression analysis based on yearly data is not useless when it comes to examining the question whether in the United States the rate of increase of productivity has been rising.

Testing the upward tendency of the rate of productivity increases (semi-logarithmic nonlinearity) with yearly data is of interest only for the period *following* 1919 (or conceivably for a period *ending* in 1919). Kendrick has shown conclusively that *if* one decides to limit oneself to semi-logarithmically *linear* relations of productivity to time (i.e., of the logarithm of produc-

therefore be placed on tendencies observable by comparing typical data for a limited number of consecutive periods, and importance will be attributed to the fact that when the Residual is used as a point of departure for estimating benefits, and the costs are defined as those of our progress-generating inputs, the benefit-trends and cost-trends show a fairly consistent relationship that "makes sense."

The method of the Residual can be employed in different ways. Kendrick has

tivity to time itself), then the slope is distinctly steeper for the period starting in the late part of this century's second decade than for the period preceding that "break." However, Kendrick did not discuss the question of the quality of the linear fit for the *pre-1919* or the *post-1919* span. The nonlinearity observable for the post-1919 is a matter of importance for the present study.

As for the period 1919-66, the upward creep of the semi-logarithmic slope, presently to be illustrated in the text with reliance on benchmark years, does express itself *also* in the statistical significance of the positive coefficient of a second-degree term. This statement concerning "significance" implies the use of yearly data, and the diagnosis based on such regressions (rather than on benchmark years) becomes somewhat complicated by the fact that if a third-degree term is also introduced, statistical significance is found for the *negative* coefficient of that term too. The inflexion point so obtained lies in the past. Yet the date of the inflexion point and inspection of the plots suggest that this results from the exceptional productivity-increases during World War II: i.e., the inflexion point reflects the slope's war-time peak which, despite the consistency of the slope's rising tendency for successive periods of longer duration, has not been reached again. One would of course not rely on such equations for projection into the future, since the future course of productivity-increases will be strongly influenced by the course of the progress-generating inputs; *but at any rate one would try to smooth out, rather than to emphasise, the effect of the depression of the 'thirties' and the effect of World War II.*

For 1919-66 I used merely forty-one yearly data for each variable (instead of forty-eight) because, as will be explained in the text, I used preliminary revised estimates for several recent years and consistency had to be observed in linking these to earlier years. It should be added that serial correlation is very high in this material. On the whole, I prefer to make use of the benchmark-year data of the text—i.e., of typical values for periods of reasonable duration—rather than of regressions computed from yearly data. The procedure used in the text is no more "arbitrary" than the decision to disregard the third-degree term in the regressions just discussed.

measured a magnitude he calls total factor productivity, this being a productivity measure from which he *excludes* the consequences of allocational shifts among sectors (his industry groupings). Hence here the shifts, which in fact have favored the higher-productivity sectors (industry groups), are made to express themselves as increases of weighted physical inputs rather than of productivity. Further, Kendrick has measured a magnitude which he calls "output per unit of total input" and which differs from his "total factor productivity" only in that it registers the results of the intersector shifts as productivity-increases rather than as increase of weighted input-quantity. Both concepts relate to the private domestic economy.

Both measures used by Kendrick register larger productivity-increases than those consistent with the usual two-factor classroom models, because his input-estimates include estimates of farm land and of real estate (site land) as part of his estimate of capital. The relative fixity of this input reduces the growth-rates of physical inputs, thus increasing the Residual. Aside from this, some technical considerations suggest that, by their own standards, the Kendrick measures of productivity-gains err somewhat on the high side but other considerations point in the opposite direction.

From 1900 to 1929, factor productivity, including the favorable effect of intergroup shifts, rose at an annual compound rate of 1.8 percent (during the last ten years of the period it rose at 2.0 percent); from 1929 to 1948 the rise occurred at the rate of 2.3 percent; from 1948 to 1966 at 2.8 percent, with no acceleration during these past eighteen years. If we exclude from the productivity-increase the favorable effect of Kendrick's intergroup shifts, we obtain somewhat smaller productivity increases but broadly similar trends for these, except that some further accelera-

tion is observed for the comparison of 1957–1966 with 1948–1957.²⁵ Kendrick's published series end with estimated values for 1957, but I made use of his somewhat revised figure for 1957 and of his figures for later years, i.e., of unpublished data he kindly let me see, though with emphasis on their preliminary character.

Thus, from the period 1929–1948 to the present, the rate of increase of productivity has risen by about 20 percent of itself; or by about 30 percent, if the effects of the intergroup shifts are excluded from the productivity-increase. From the opening decades of the century to the present, the rate of productivity-increase has risen by about 60 percent (a figure which becomes 75 percent with exclusion of the intergroup shifts).²⁶

Indications of a rising tendency of the rate of productivity-increase are consistent, though the rise is small as compared to the rise of the recorded progress-generating inputs, particularly of *R & D*, in relation to output. The discrepancy becomes significantly reduced by reasonable allowance for progress-inputs that are not recorded as such.

IV. *The Framework for Productivity and Benefit Measurement: How Serious are its Pitfalls?*

1. *Using the All-Inclusive Cost-Base Requires a Deduction on the Productivity-Side.*

We start with a pitfall for the avoidance

²⁵ On this basis the following results are obtained: from 1900 to 1929 the rise was 1.5 percent (for the sub-period 1919–29: 2.0 percent); from 1929 to 1948 it was 2.0 percent; from 1948–57 it was 2.3 percent; and for the following nine years 2.6 percent.

²⁶ Denison, who has made an attempt to separate quantitatively the effect of different sources of growth—an attempt which because of interactions between his “sources” I am inclined to consider too ambitious—used his estimates *also* for explaining what results he would obtain by a method much more similar to that of Kendrick and of other investigators. It turns out that from the period 1909–29 to 1929–47 he would obtain a 45 percent rise of the rate of productivity-increase. See Edward F. Denison, pp. 148–49.

of which a definite prescription can be given. If we say, as we do when using our all-inclusive cost-base, that productivity-increases cost us *inter alia* the inputs needed for increased per capita education, then we are implying that on the benefit-side the bulk of this particular cost must be deducted from the productivity-increases. The reason is that a once-attained higher level of per capita education must subsequently be maintained to keep the level of knowledge from becoming reduced. In this regard the characteristics of education are different from those of *R & D* or from those of the costs of changeovers.

What needs to be deducted on the benefit-side is the “bulk” of the cost that had been entered for education—not the entire cost—because we are relating our magnitudes to output, and constancy per unit of the population means a gradual reduction per unit of output. However, on reasonable quantitative assumptions the “discounted value” of the difference between these two growth rates is small, and hence I will indeed deduct the *bulk*.²⁷

2. *The Embodiment-Disembodiment Issue, the Quality-Quantity Controversy and the Importance of “Dimensions of Improvement.”*

We shall now consider the main problem of the present section. Our reasoning can be placed into frameworks of different

²⁷ If progress were limited to a brief interval, then not merely the “bulk” but, for practical purposes, *all* of the cost-item entered for increased per-capita education would have to be deducted on the benefit-side of the ledger. The increase in per capita output resulting *merely* from, say, a 2 percent yearly increase of the capital-labor ratio “at a given level of technology” would require a negligible correction of the kind described in the text. But if we consider a very long sequence of periods with technological progress and assume, say a 2 percent yearly increase in per capita output then, with a 10 percent discount rate, roughly four-fifths (or a shade less) of the cost-item entered for increased per capita education should be deducted on the benefit-side. I will in fact deduct about four-fifths, with some rounding where this is called for (see fn. 40, for the calculation of the 1966 and the 1953 yield).

kinds as long as these have enough in common with the models discussed in Appendix A. For us the essential property of these constructs is that they are models of disembodied progress in which the elasticities of output with respect to the inputs are reasonably well approximated by the distributive shares of the inputs, and in which distributive shares remain reasonably stable. To the extent that these assumptions are unrealistic our analysis needs to be adjusted, and in Section III, I have already pointed out that one type of adjustment does not change the conclusions very much. What should we think of the "disembodiment" implication of our framework, i.e., of the implication that progress leads *inter alia* to the production of new types of goods rather than *depends* on new equipment?

Models of so-called *embodied* progress do not seem convincing to me on logical grounds, and to my knowledge, no claim of empirical superiority has been made, let alone established, for these frameworks. As I said, all progress is necessarily disembodied *in the sense* that new ideas must always be put into effect with reliance on the *initially given* resources. This is an essential constraint under which all economies operate. Improved production with the initially given resources then leads to more and better capital goods—hence to the replacement of old with superior equipment and structures ("obsolescence")—and it yields more and better consumer goods. But all this represents *forward*, not backward, embodiment.

For various purposes it has proved rewarding to interpret progress as "factor-augmenting"; yet even if progress is interpreted as purely *capital*-augmenting, this alone does not change the framework into one of embodied progress. Using a framework in which some input is "augmented" merely means expressing the productivity-increase of that input as the equivalent of a

specific quantity-increase *and* expressing the productivity-increases of other inputs as resulting from complementarity. No empirical superiority has been claimed for this degree of unilateral emphasis on capital—i.e., for assuming pure capital-augmentation—but at any rate more is needed for obtaining "embodiment," in a distinctive sense of the term.

If the embodiment hypothesis is not to lose what I consider its essential characteristic, the additional assumption must be made that each period's capital-goods output can become subject merely to a single dose of augmentation which describes a specific level of technology; and that only the next period's capital goods—the next vintage—can become subject to the next dose. This, I submit, is an unconvincing assumption if we take into account *why* next period's capital goods—next period's equipment-models—are superior to this period's equipment-models. They are superior because this period's *initially given* equipment-models and labor perform better than how these inputs would have performed without the new ideas. So why should the possibilities with a period's capital goods—a given vintage—become *exhausted* during a specific period for which a specific level of technological knowledge is defined? If the possibilities do not become exhausted in this fashion, then a level of technological knowledge ceases to be uniquely described by (or embodied into) a specific vintage of capital goods.

In the absence of strong empirical support for the hypothesis of pure capital-augmentation and of the required additional assumption, I have fewer misgivings about disembodiment than about embodiment. Yet while so-called embodiment makes too much of the role of new physical capital formation in bringing about progress, disembodiment implies overlooking its progress-promoting effect which probably does exist. The existence of this ef-

fect complicates matters, and I shall soon return to this difficulty.

We must be mindful also of another property of the framework here employed. We are implying, as does the method of the Residual in general, that it is advisable to try to translate quality-improvements of consumer goods into additional quantities and yet to express productivity-increases of capital goods and of inputs in general as quality-improvements for given physical quantities.

In my appraisal there is no reason to become disturbed about this lack of symmetry between the interpretation of improvements of inputs and that of improvements of consumer goods. The alternative would be that of lumping together our "progress-generating inputs" with the other inputs, and of interpreting the research worker not as one who *increases* the productivity, and hence the quality of the inputs with which measured output is produced, but as one who *has* the same kind of productivity in the production of measured output as do (say) production workers.²⁸ However, for my present purpose it would be wrong to cut the story short in this fashion. I am concerned with the question how the social yield of progress-inputs compares with that of other inputs, and this concern calls for placing emphasis on the indirectness with which the progress-inputs produce additional measured output.²⁹ Thus, for the present purpose it is useful to have a framework in which the progress-inputs increase the "quality" of the other inputs and in which the other inputs produce quantities of output.

In practice the effort to let the produc-

tivity-increases of inputs express themselves in quality-changes for given physical quantities, and to translate the quality-improvements of consumer goods into quantities, gives rise to many procedural difficulties and the methods for taking care of these are admittedly very imperfect. In particular, it has been argued that the quantitative allowances for the improvement of consumer-good quality are not only very crude (which they certainly are), but in all probability also insufficient. In a general way it makes good sense to suspect that the allowances are quantitatively insufficient, but we should nevertheless note that the standard of living of Western countries is rising rapidly, and it is particularly true at higher levels of income that the market basket illustrates instances of consistent quality-deterioration as well as of improvement. Nevertheless, for consumer goods I shall make an allowance for the assumed underestimate of net quality improvement (see Appendix C).

Last but not least, an inevitable imperfection of frameworks of the kind here employed results from two-way interactions between physical capital formation and technological progress. As I said above, my misgivings about the method by which vintage models deal with this relationship does not mean that I consider it harmless to disregard the interactions.

Starting with initially given resources, the chances of substantial and continued improvement are undoubtedly greater if the processes used are *rich in changeable properties*—i.e., in what I elsewhere called *dimensions of improvement*—than if these processes offer improvement possibilities only in a few directions; and *the use of a large variety of instruments of production opens up a large number of dimensions of potential improvement*. In another study I found strong indications that the results of learning-by-doing are significantly influenced by differences in the available

²⁸ See Zvi Griliches and Dale Jorgensen.

²⁹ Looking at the matter in this fashion is of course compatible with using models that represent progress as the "augmentation" of inputs. Indeed, models of this sort usually draw a sharp distinction between "inputs measured in physical units" and "inputs measured in (augmented) efficiency units."

number of dimensions of improvement.³⁰ So physical capital formation helps the improvement process, if for no other reason than because it increases the number of directions in which improvement can proceed. The interaction is mutual: we have seen that without *any* technological progress whatever, net physical capital formation could not exceed the equivalent of the growth of other inputs because less labor-intensive machines are different machines. Some readers might therefore prefer the alternative framework outlined at the outset of Section III (see also fn. 21).

On the whole, I do not believe that we would become guilty of gross error by disregarding these mutual interactions when using our framework for distinguishing between progress-generating activities and physical investment *provided* we stay aware of the fact that large changes in the scale on which either of these activities is pursued have implications for the conditions under which it is possible to engage in the other type of activity. Of this we must remain aware.

V. Progress-Generating Inputs and Costs: The Significance of the "Unrecorded" Items³¹

From 1953 to 1966 total recorded *R & D* rose from \$5.2 billion to \$22.2 billion, and this represents a better than twofold increase when *R & D* is expressed as a proportion of the privately produced

GNP. This proportion rose from 1.6 percent to 3.3 percent.³² While an appreciable part of this increase was "real"—i.e., "physical"—considerable importance attaches to changes in relative values.

The question of the weight of the *real-input component* versus the *relative-valuation component* of the *increase* of the ratio of *R & D*-cost to private *GNP* is discussed quantitatively in Appendix B. This question is of some importance because the relative-valuation component represents in part rent-formation—a mere transfer of income—rather than a rise of social costs. However, in any event, *even the relative-valuation component* represents in good part rising social costs, because the relatively-rising salaries of research workers (and also of teachers) have attracted into the progress activities individuals whose qualifications and productivity would be higher also in other occupations. This must have been the driving force of the transfer. Such a phenomenon is accompanied by rent-formation, but I suspect that in the progress-activities—particularly in research and higher education—the rent-formation accompanying the rise of social costs is smaller than in the types of activity mostly referred to when the nature of the transfer-rent problem is illustrated.³³

³² In 1953 the private *GNP* was \$332.7 billions (the total *GNP* \$364.6 billions), in 1966 the private *GNP* was \$671.1 billions (the total \$747.6 billions).

Kendrick's productivity increases are measured merely for the private *domestic* output, but when expressed in terms of proportions the difference between "private domestic" and "private" is exceedingly small. Also, for recent periods there is a strong presumption that those parts of the foreign economies from which domestic income is derived have greatly benefited from our domestic progress-generating activities, and that these foreign sectors have shown a productivity-increase at least as great as that of the corresponding domestic sectors.

³³ In the typical illustrations, transfers from one sector to another create rents because, given a homogeneous labor force in the favored sectors, *all* workers of these sectors must be paid the higher wages which only *some* could earn in the occupations from which workers are becoming diverted. I conjecture that where the pro-

³⁰ See my "Specific Interpretations of Learning by Doing."

³¹ The numerical statements in this section will be based on National Science Foundation (*NSF*) data for recorded *R & D*, and on data of the Department of Health, Education and Welfare (*HEW*) for education expenditures, except that Fritz Machlup's book, to which reference will be made in the section, was used for broadening the concept of education costs much beyond the *HEW* coverage.

The reader will find the relevant *NSF* data in the periodically appearing *NSF* publications listed in the Reference section.

See References also for the *HEW* data.

We have just seen that from 1953 to 1966 the ratio of total recorded *R & D* cost to private *GNP* rose by more than 100 percent. For the period 1929–1953 only sporadic and incomplete data are available, but there is good reason to conclude that during this period there occurred a very significant increase of *R & D* inputs in relation to real private *GNP*, and that in terms of undeflated values the increase was even more substantial than in real terms (see Appendix B). These increases are out of line with the rise of the rate of productivity-increase. The latter is a 60 percent or, at most, a 75 percent increase *even* if we go back to 1900–1929 and thus compare the *pre*-1929 period with the present.³⁴

The efficiency of the *educational complement* of research would to some extent become adjusted to the requirements set by new knowledge-production even if per capita knowledge inputs did not rise in quantity. In reality the quantity of these inputs has risen appreciably. Partly from *HEW* data and partly from those of Fritz Machlup, I infer that from the calendar year 1953 to 1966 the cost of increased per capita education rose from the equivalent of 0.3 percent of private *GNP* to the equivalent of about 0.5 percent (see Appendix B). Increased education works itself out with significant lags, but an appreciable rate of increase dates back far into the past. Therefore, trend-appraisals are not appreciably distorted if we associate 0.3 percent with the 1953 benefit, and 0.5 percent with the 1966 benefit, though the levels of the rates of progress-return *are* somewhat distorted—i.e., are calcu-

lated as if the lag did not exist—a fact that need not disturb us much because the smaller cost of the past, which in view of the lag would represent a more appropriate present charge, arose correspondingly earlier and for that reason should be viewed as the equivalent of more than the undiscounted figure suggests.³⁵

As for the industry-financed (institutionally profit-oriented) *R & D*, this rose from \$2.2 billions in 1953 to \$7.2 billions in 1966. When expressed in relation to private *GNP* we obtain a rise from 0.7 percent to 1.1 percent. In Appendix B it is shown that the *post*-1953 rise of *this* particular proportion is likely to have been mainly (perhaps entirely) the result of the increase in the relative prices of *R & D* inputs. That is to say, here the relative-valuation component accounts for practically the entire effect, while for *total R & D*, the real-input component also was significant (even for 1953–1966).

Regardless of whether we are interested in yields on our all-inclusive cost-base or in yields on a cost-base limited to institutionally yield-oriented progress-inputs, exclusive concern on the cost-side with recorded *R & D* and with increased per capita education would lead to the twin conclusions that yields have been declining steeply and that they nevertheless still are at a “perplexing” level. Such calculations are misleading because they disregard the fact that obsolescence in the conventional sense, as well as some costs of a character similar to that of obsolescence, and also some costs of a different kind, are omitted from the *recorded* progress-generating costs. This is so even if for the present purpose we regard the education costs in Machlup’s broadened sense as “recorded.” Many of the *unrecorded* costs, including obsolescence itself, fall in the category which in Section I, I called costs of changeovers to

gress-sectors are the favored ones the significance of this phenomenon is reduced by the fact that the labor force of research establishments is far from homogeneous. The reason why a high proportion of these workers is facing increasingly favorable market conditions is that a high proportion is gradually improving its qualifications which are getting to be considered increasingly valuable also in other occupations.

³⁴ See the concluding part of Section III.

³⁵ The appropriate discount rate is obviously in excess of the growth rate of the output base.

new methods. Their weight is substantial, but for a given rate of technological progress these costs are unlikely to change much in relation to private *GNP*. Hence allowances for these costs reduce not merely the rates of return but *also* their downtrend.

It is easy to become misled into the belief that the method of the Residual automatically implies proper recognition of obsolescence. In a sense it does recognize obsolescence but it does not do so in the sense relevant to our problem. The method is not intended to, and it does not, take account of the fact that in any of the "base periods" with which the "present period" is compared, the *now old technology was new*.³⁶ Hence at that time the *now old* technology gave rise to obsolescence and it required various *instrumental* activities that represented costs of changing over to new methods. These costs would not have been incurred in the base-period and they would not be incurred now if there had been no technological advance in either period. In a continuously progressing economy the inputs used for the obsolescence-component of capital-replacement are continuously withheld from the production of consumer goods or from net investment (or from the public-sector counterparts of these).³⁷

A list of the items that may be regarded as unrecorded costs is presented in Appendix C, along with explanations. In spite of the large number of items on the list, we may take it for granted that only one item—conventional obsolescence—corresponds individually to an appreciable pro-

portion of private *GNP*. As for the other items, in some cases it is doubtful whether they have a legitimate place on the list and at any rate, even their joint weight cannot be appreciable when expressed in relation to private *GNP*. Among the items listed in Appendix C there are only a few that would be suspect, even at first sight, of possessing large weight. Item 5 of the Appendix—the cessation of learning-by-doing with the old equipment and for the old product when new methods are introduced—belongs among the suspects, but in reality it is unlikely to be of great quantitative importance because in each period the discarded methods are apt to be those for which learning-by-doing has largely run its course. Item 6—the technical imperfections of recording—must have been large in the early decades, but we are concerned here mainly with the post-1953 period, and our few references to pre-1953 periods were based on attempts to get indirectly at what *now* would be "recorded" *R & D*. For the present purpose I shall assume that a deliberately high estimate is obtained of the joint weight of the unrecorded costs *other than obsolescence* in the usual sense if these *other* costs are put at 2 percent in relation to private *GNP*; and that a deliberately low estimate is obtained by putting these items at 1 percent.

What *mainly* matters is obsolescence in the usual sense. The fact that we consider a component of capital consumption a cost even though we are moving in a *GNP* framework may seem paradoxical, but for no good reason, because while for our periods *GNP* may safely be used as a proxy for *NNP*, we must take into account that if capital consumption had all along been smaller, then any given *GNP* would have been a proxy for a higher *NNP*.

Experts have repeatedly expressed the view that practically all of the capital consumption and replacement, which cor-

³⁶ To be precise, the method "nets out" obsolescence automatically *only* to the extent that, from the base period to the period with which it is compared, obsolescence *may* have *increased* in relation to the output-base.

³⁷ The same is true also of obsolescence in the public sector. This is perhaps best regarded as one of the social costs of progress in the private sector (but the quantitative importance of how this item is interpreted is very small for our purpose).

responds to more than 10 percent of *GNP*,³⁸ reflects obsolescence, i.e., a result of technological progress. Such a statement needs to be reinterpreted for our purpose. If the service lives of capital goods were to be prolonged significantly, then, unless the equipment-goods and the construction industry were to raise their costs and selling prices for the sake of durability, the users of capital goods would be spending more on repairs and maintenance. The question here is: how much more? This, I believe, has remained an unexplored question which would deserve a research effort. In the absence of dependable information, I suggest two guidelines for the present purpose. I will base my subsequent reasoning more on the second than on the first of these.

1) By assuming that total capital consumption typically amounts to about 12 percent as expressed in relation to *private GNP*, and by charging one-fourth of this to technological advance for obsolescence, one is unlikely to overcharge the progress-

³⁸ From 1919 to 1955 Kuznets' capital consumption allowances move in the range from 10.9 to 15.1 in relation to *GNP* (constant prices); at the beginning and also at the end of the period the figure was almost precisely in the middle of the range. This implies Kuznets' Variant III of output. Using other Variants of Kuznets, wider ranges are found for movements of the depreciation ratio in the region *above* 10 percent. See Kuznets' Appendix A, Table R-2, p. 487.

The capital-consumption ratio implied in the *Capital Stock Study* of the U.S. Department of Commerce, Office of Business Economics seems to be somewhat smaller than Kuznets', but also typically somewhat higher than 10 percent. See for this e.g. Robert C. Wasson's articles in the *Survey of Current Business*, particularly in the December 1966 and in the February 1969 issue. However, proper interpretation of the relationships here considered requires also use of computer output available from Mr. Wasson's office. My evaluation implies a reasonable "blowing up" of the Commerce Capital Stock Study depreciation figures, in view of the fact that the figures of the Study relate only to the business sector including farms, i.e., exclude the residential sector as well as general government. I was using the Capital Stock Study's straight-line depreciation estimates and was using them on "Cost 1" basis. This cost-basis makes the method of deflating comparable to that applied to the Commerce estimates of the *GNP*.

account. One way of visualizing the implications of a charge of such size—3 percent of private *GNP*—is to note that it is consistent with the hypothesis that in order to arrive at a doubling of the useful service lives, the users of capital goods would have to increase their present expenditures on current repairs and maintenance by the equivalent of about 50 percent of what would then become their reduced current capital-consumption cost.³⁹ For structures which represent a very weighty component of the capital stock, this is quite likely to be an overstatement of the costs of extending the now usual service lives, hence an understatement of the costs to be set against technological advance and against the rise of living standards in which it expresses itself. The same may be true of *some* types of equipment, but there undoubtedly exist types for which these numerical assumptions understate the costs of the extension of service life, and thus overstate the cost of progress. On the whole, charging 3 percent of private *GNP* is unlikely to err (or to err *much*) on the high side of costs, and this charge might well err on the low side. As long as we try not to overcharge we may add merely 1 percent of the private *GNP* for unrecorded costs other than obsolescence, i.e., for the *other* costs listed in Appendix C. We arrive at 4 percent of private *GNP* for what I consider a "moderate" estimate of the unrecorded costs, though I will not build much on this figure.

2) An immoderate overstatement of the obsolescence-cost of progress would be obtained by charging to progress the *entire* capital-consumption—i.e., about 12 percent of private *GNP*—and by adding 2

³⁹ Any hypothesis of this kind has the realistic implication that in a progressive economy the "representative" physical asset is produced in such a way as to be in some sense "too durable." This has several reasons, one of which presumably is that lack of durability is not the only unfavorable property of a "flimsy object."

percent for the other unrecorded costs, thus obtaining a charge corresponding to 14 percent of private *GNP*. But it will prove to be important to take a look also at the progress-yields obtained on this immoderate assumption. I *will* build on the fact that the 14 percent charge is excessive by a significant margin.

VI. Levels and Trends of "Objective" Yields and Observations on Matters Calling for Subjective Appraisal

The method by which the quantitative yield-appraisals were obtained is based on the discussion in Sections III and V and the details are explained in the footnote below.⁴⁰ Note that on the productiv-

⁴⁰ (a) 1966, all-inclusive cost-base, 4 percent of private *GNP* charged for unrecorded costs. Recorded total *R&D*, as a proportion of private *GNP*: 3.3 percent (all other data will also be expressed in relation to this output measure). Cost of increased per capita education: 0.5 percent. Annual rate of productivity-increase: 2.8 percent, which becomes reduced to 2.4 percent when the bulk of the foregoing education item is deducted (see Section IV, 1). Since $3.8 + 4.0 = 7.8$, and since $2.4 \div 7.8 = 0.31$, the resulting rate of return is 31 percent. For 1966 this rate and that presented in (c) below, bound the range defined in Section II when the charge is 4 percent.

(b) Same as (a) except with the clearly excessive charge of 14 percent instead of 4 percent (see last paragraph of Section V). The resulting rate of return is 13 percent. For 1966 this rate and that presented in (d) bound the range defined in Section II when the charge is 14 percent.

(c) 1966, reduced cost-base, with 4 percent charge. We enter 1.1 percent for the recorded *R&D*. Since $1.1 + 4.0 = 5.1$, and since $2.8 \div 5.1 = 0.55$, the resulting rate of return is 55 percent. See last sentence in (a).

(d) Same as (c) except with 14 percent charge. The resulting rate is 18 percent. See the last sentence in (b).

(e) 1953, all inclusive cost-base, with 4 percent charge. Recorded total *R&D*: 1.6 percent. Cost of increased per capita education: 0.3 percent. Productivity-increase 2.8 percent which becomes reduced to 2.55 percent when the bulk of the foregoing education item is deducted. Since $1.9 + 4 = 5.9$, and since $2.55 \div 5.9 = 0.43$, the resulting rate of return is 43 percent. For 1953 this rate and that presented in (g) bound the range defined in Section II when the charge is 4 percent.

(f) Same as (e), except with 14 percent charge. The resulting rate of return is 16 percent. For 1953 this rate and that presented in (h) bound the range defined in Section II when the charge is 14 percent.

(g) 1953, reduced cost-base, with 4 percent charge. The

ity-side we rely on Kendrick's estimate of annual productivity-increase *including* the effect of intergroup shifts (2.8 percent), but the reader can convince himself that the result would be influenced little by the use of estimates that exclude the intersector effect (for the most recent period this would have been an increase of 2.6 percent p.a.). To me it seems reasonable to relate also the intersector effect to progress-generating inputs.⁴¹

Many statements one would like to be able to make about these results would be much too risky for presentation. But there are two statements which I consider to be safe.

The first of these is that the present average social rate of return from the progress-activities is *substantially* in excess of 13 percent on the all-inclusive cost-base, and *substantially* in excess of 18 percent on the cost-base limited to institutionally profit-oriented progress-inputs ("reduced" cost-base). The rates of 13 percent and of 18 percent would imply charging all capital consumption to progress, as if extension of service-life were entirely *costless ad infinitum* except for continued maintenance costs at the *now* usual level. If we charged one-fourth of the capital consumption to progress, and made a moderate charge also for other unrecorded costs, we would obtain a 31 percent rate of return on the all-inclusive cost-base, and a 55 percent rate on the reduced cost-base. These particular charges rest on

resulting rate of return is 60 percent. See last sentence in (e).

(h) Same as (g), except with 14 percent charge. The resulting rate of return is 19 percent. See last sentence in (f).

The lowest rate in this listing is that in (b) above. As for the "corresponding" rate obtained in the *alternative framework outlined at the outset of Section III*, see fn. 21.

⁴¹ By using the 2.6 percent figure we would reduce the lowest of the rates listed in the preceding footnote (i.e., the rate there presented in paragraph (b)) from 13 percent to 12 percent.

guesswork (I believe of a *plausible* kind), but it may be regarded as a fact that 13 percent and 18 percent understate the average rate of progress-return by a substantial margin and this will prove significant for the present analysis.

The second statement I consider to be reasonably safe is that the rates of return show some degree of downward tendency in the long run (see here also Appendix B). The qualifications one might want to add to this statement are not sufficiently persuasive to arouse much suspicion in the diagnosis that such a tendency is observable, though I would not want to assert that my figures give an accurate idea of the *extent* to which a downtrend has manifested itself. On what I defined as the reduced cost-base this downward tendency has been very mild, but it would not entirely disappear on *any* cost-base that to me seems reasonably chosen.

The suggestion of a long-run downward tendency of the average rate of return is not misleading, because the long-run increase of the costs of progress-inputs has *all along* exceeded the rise of the productivity-gains when both are expressed in relation to the same output-base. In particular, the approximately constant uptrend of productivity during the post-war decades has been associated with increasing progress-costs even if we consider merely the costs of the privately financed progress-inputs. The second half of the present decade does not fit well into the pattern, but it is too early to try to interpret the record of the very recent inflationary years, not only because the rate of resource-utilization in the American economy has risen suddenly but also because the rate of increase of the progress-inputs has been tapering off. This tapering off occurred partly because the rate of increase of space research tapered off and then space research reached a peak in 1966; and partly because even more recently

there also took place anti-inflationary downward revisions of government programs (which is not to say that the rate of increase of *R & D* might not have declined even aside from these circumstances). Consequently it is too early to ask whether in the mid-sixties the downward tendency of our benefit-cost ratios was or was not interrupted, but further productivity-increases would have to assume a very unlikely course to invalidate the conclusion that there did occur some degree of long-run decline, even with reasonable allowances for lags.

We should note that the downward tendency of the progress-yields would be slightly smaller than our figures suggest if we assumed a small rise of the rate of increase of factor productivity from 1953 to 1966. In fact there is valid reason for assuming a *small* acceleration of progress but, as concerns the downward tendency of the yields, allowance for this would make little difference, because such a modification would call at the same time for a small increase of our charge for the unrecorded costs.⁴²

⁴² I assumed a 2.8 percent rate of increase of factor productivity for the entire period 1948-66. This is in accordance with the Kendrick estimates when the favorable productivity-effects of intersector (intergroup) shifts are included. On this basis there was no increase from 1948-57 to 1957-66. However, if the effect of these shifts is excluded, then from 1948-57 to 1957-66 a rise is observed from a rate of increase of 2.3 percent to a rate of 2.6 percent. Furthermore, even the first of these two types of measure of productivity-increase rose from the period 1929-48 to 1948-66 (it rose from 2.6 percent to 2.8 percent), and it may be argued that this rise has resulted in part from the post-1953 increase in progress-inputs.

When comparing 1966 with 1953 I took no account of any such increase on the benefit-side of the ledger. I also used the same percentage allowance in relation to private *GNP* for the unrecorded costs of 1966 as for those of 1953, since there exists a general presumption that these costs tend to move in proportion to the rate of progress. During these particular thirteen years there may conceivably have occurred a somewhat more than proportionate increase in specific costs appearing in the list of Appendix C. One may be led to suspect this for example because of the increase of the white-collar proportion of the labor force. On the other hand, the

A downtrend of rates of return in the progress-generating activities is, of course, not the same phenomenon as the previously discussed gap between the average and the marginal rate on the progress-inputs of single periods; the downtrend over time calls for a different interpretation. It may be interpreted as meaning that researchers, teachers and other workers engaged in knowledge-acquisition and distribution activities of *growing complexity* have increased the efficiency of their own group *somewhat less* than would have been required for unchanging performance at unchanging costs. Even the relative-valuation (or relative-wage) component of a downtrend can in good part be so interpreted, because while this component becomes enlarged by rent-formation, it nevertheless is true that the relative increase of the incomes of progress-workers typically results from the need to use very highly qualified personnel on a rising scale.⁴³ Trying to avoid a downtrend of the rates of return on these particular inputs by limiting the scale of the progress-activities of each successive period accordingly would in the long run prove incompatible with a nondiminishing rate of technological advance in the economy at large. In view of the scarcity of the ultimate re-

sources needed for any economic process—including the process of self-improvement of progress-workers—a secular downward tendency of yields should indeed be expected, and Kenneth Arrow was right in reminding us at last year's annual meeting that "eternal exponential technological growth is just as unreasonable as eternal exponential population growth."⁴⁴ But it is equally true and important that such a slow, secular downward tendency of the progress-yields may become interrupted for *very* long periods. Recurring "break-throughs" play an important role in the history of the sciences.

Quite aside from the question of a downward tendency with the actually observed composition of the progress-activities, we should recall that, given the present scale of these activities, even a person in full agreement with the community-decisions concerning composition would in all probability have to consider the *marginal* social rates of return of each period lower than our average rates. This gap arises partly because high-quality personnel is limited, and partly because the solution of problems presupposes the solution of earlier ones. The least promising project of any period is apt to be appreciably less promising than the period's average project. But as I said in Section II, I am abstaining here from efforts to apply formal analysis to the problem of marginal rates, since each individual is entitled to his own marginal calculus and in most of the area in which this paper has been moving individual valuations become submerged in compromise rather than become validated at the margin by means of purchase, sale and production.⁴⁵ This is why I have so far

weight of the unrecorded research inputs listed in (6) of the same Appendix is likely to have declined.

⁴³ From 1953 to 1966 the full-time-equivalent number of *R & D* scientists and engineers rose from about 225,000 to about 520,000, and as a proportion of civilian employment it rose from four-tenths of one percent to seven-tenths of one percent. At the same time the salaries of the *R & D* scientists and engineers rose appreciably in relation to wages and salaries in the economy at large. The sample-data available on this change in the wage-salary structure point to a relative salary-increase of roughly 20 percent in favor of *R & D* scientists and engineers. This does not appear to be very different from the relative salary change in favor of scientists and engineers in general, including those employed in *R & D*; proportionate representation of scientists and engineers in the employed labor force rose during our period by slightly more than 50 percent of their 1953 proportion and in 1966 their number was about 1,500,000, i.e., about 3 times that of the *R & D* scientists and engineers.

⁴⁴ See Kenneth J. Arrow. See also my study of learning by doing, August 1969.

⁴⁵ As was pointed out before, the teaching sector does offer some room for such marginal equilization, but only in a very limited way because even where a price is charged, resource-use by teaching institutions results very largely from political or quasi-political decisions. See second paragraph of Section II.

914414

focused on average rates of return and on "necessary conditions," using what I believe to be very widely shared assumptions as to the meaning of *low* and of *high* valuation of controversial nonmarket contributions.

At this point, attention should be directed at those aspects of our problem which call for highly subjective judgment. My illustrations will be found in Appendix D, which contains observations on significant structural and allocational problems within the area of *R & D*. Prominent among these problems is that of "concentration" in several senses of this term. As for educational policy, it has become all too obvious that the future of Western civilization depends on the personal attitudes millions of individuals will develop to structural problems belonging in that sector of the progress-generating activities.

We may now return to our point of departure. As long as we want to remain on a level of reasonable objectivity, we can only try to play into the hands of those who must make their own value judgments. On alternative cost-bases it is possible to make reasonable statements of fairly general validity on the social profitability of the progress-generating activities as a whole and on the trends of this profitability. For the time being the average social yields are high on both our cost-bases, *very much* higher than 13 percent and 18 percent, respectively, as was seen. It should of course not be overlooked that a waiting period of considerable duration elapses between *some* progress-generating inputs and the onset of their yield. But this is true only of *part* of the progress-inputs as we have defined them. Furthermore, when comparing these yields with yields from physical capital formation, another lag-adjustment is also needed, and as explained in Appendix A, this works in the opposite direction (in favor of the progress-inputs). We should remember also that

while on the one hand the lag between progress-inputs and their results would call for *upward* revaluation of the inputs in relation to their results, on the other hand, earlier progress-inputs have all along been smaller progress-inputs.

At any event, regardless of what lag-adjustment we make within reason, the average social rate of return—even the rate on the all-inclusive cost-base—satisfies our necessary condition with ease. This "real rate" is at present much higher than the marginal rate from physical investment at a more or less given level of knowledge. The latter rate should be estimated at less than 10 percent; a good case can be made for estimating it at a figure located in the range between 5 percent and 10 percent.⁴⁶ Moreover, the rates on progress-inputs suggested by our analysis are appreciably higher also than the typical pre-tax corporate profit-rates, a fact which is worth noting even though no comparability is claimed for our rates of return with accounting profits.⁴⁷

⁴⁶ A rough estimate of the rate in question, which is obtained by way of multiplying the ratio of yearly output to the capital stock by the share of reproducible capital in income, points to less than 10 percent. See Section II.

⁴⁷ The published corporate profit rates (which are related to the book value of equity) are of course private rather than social rates. This would not make too much difference if the private rates were rates "at a given level of technological knowledge" but in reality they are earned *partly* on private progress-generating inputs which do not enter into the book value. Also, in contrast to our "real" rates these corporate rates are money rates earned in large part on equity such as has been paid up or accumulated in much earlier years than those in which the profits were earned. There are several further reasons for noncomparability. On balance there exists a *strong* presumption that the corporate rates on book value are *considerably higher* than the real rates on physical investment. The published pre-tax profit rate on the book value of American manufacturing corporations as a whole has of late been about 20 percent, and even this is less than the rates of progress-return to which our calculations in the text point. Again with no claim to comparability, it may be added that the earnings-rate on the *market value* of Standard and Poor's common-stock sample (not wholly limited to manufacturing) is in the neighborhood of 6 percent,

For an appreciable period to come, even a slow downtrend of the progress-yields would leave a significant differential between the average social rate of return from the progress-activities and the conventionally defined marginal rate of return from physical investment. One is tempted to add that, given the size of this differential, it should be possible to promote the gradual shifting of a larger proportion of our inputs into the various progress-activities without forcing the *marginal* social rates on these below the marginal rates on physical investment (the latter being an activity which does not come in fixed proportions with technological progress). In my appraisal, this statement, too, makes very good sense. But even if most people should agree with this statement, and yet there should be substantial disagreement on the nature of the projects that satisfy the marginal criteria, the possibility of increasing the weight of the new-knowledge sector would still depend on how well the political mechanism is capable of bridging the differences. In view of this, I will end with a question in bargaining theory: Is it realistic to expect that the propensity to reach compromises can be increased by making the bargaining parties aware of the fact that the joint payoff on reaching agreement is high?

APPENDIX A

Relevant Properties of the Models Referred to in Section I

In the Cobb-Douglas model neutral disembodied progress⁴⁸ expresses itself in an increase of the so-called Cobb-Douglas constant; hence in a growing economy the absolute contribution of once-achieved progress will in the future be stepped up in

proportion to increases of the output-base itself. As expressed in proportion to the output-base, the progress-generated increment will be a permanent flow at a constant rate.

From a rate-of-return concept for knowledge-acquisition, one expects comparability in this regard with net output-increments attributable to permanently maintained additions to the physical capital stock "at a given level of knowledge." To the extent that increased labor inputs will be responsible for the future rise of the output-base in relation to which the progress-generated net addition to output develops, the Cobb-Douglas framework does indeed assure comparability with physical investment in this regard. Not only the absolute contribution of present technological progress *but also that of present physical investment* at a given level of knowledge will be raised in proportion to output, because in the Cobb-Douglas framework the growth of the labor input (given the input of capital) raises total output and the average and the marginal product of capital in identical proportions.

On the other hand, to the extent that in a growing economy further capital formation will be responsible for the future rise of the output-base, there does not exist the same comparability of the progress-generating effort with physical investment. To this extent any observed numerical ratio of *progress-generated* output-increments to the present cost of progress corresponds to a higher social yield than does the same numerical ratio of *investment-generated* output-increments to the cost of investment, provided we express the output-increments as well as the costs in proportion to the present output. This is because future physical investment will not increase the absolute output-increment generated by present physical investment while future physical investment *will* increase the absolute output-increment developing as a result of present technological progress. Yet on any realistic assumption concerning the relevant discount rates and the long-run trends in the supply of savings for capital formation, the resulting difference or non-comparability is small and it may be interpreted as the equivalent of a small post-

which is of course after corporate income tax and would correspond to a shade more than 10 percent before that tax.

⁴⁸ For the disembodiment implications, see Section IV, 2.

ponement of the onset of the yield of the progress-generating activities.⁴⁹ We will have to accept this noncomparability. This difference is probably outweighed by another difference which our lagless approach will at first also neglect but for which we will have to make crude allowances subsequently. I mean the length of the gestation period elapsing between *some* of the progress-generating inputs and their industrial results, as compared to the shorter gestation period for physical capital formation at a given level of technological knowledge.

Our conclusions for the Cobb-Douglas model are valid also in the CES framework, provided we supplement the latter with a distributive-shares equilibrium mechanism and assume that the economy is moving in the neighborhood of that equilibrium (with pure labor-augmentation).⁵⁰ In such a model, given degrees of now achieved labor-augmentation will in all future have to be multiplied by the same weight to give the proportionate rate of progress-generated net addition to output, and hence the proportionate output-increment will remain constant, as a permanent flow. From here on—i.e., concerning the comparison with physical investment—the reasoning is the same as for the dynamic Cobb-Douglas model.

For the relationship between these models and “reality,” see my December 1967 article in this *Review* to which I referred in Section I. See also Sections II and IV.

APPENDIX B

(see Section V)

Explanation of Statements on

Input-Trends

1) The “Real-input” Versus the “Relative-value” Component of the Rise of Recorded *R & D*.

⁴⁹ On the generous assumption of a 1 percent per annum increase of the output-base as a specific result of physical investment, and assuming a 10 percent discount rate, this noncomparability is the equivalent of a postponement of the onset of the yield of progress by roughly one year.

⁵⁰ In this framework the elasticity of substitution is smaller than one and hence labor-augmenting innovations are also *labor-saving in Hicks's sense*, though in shares-equilibrium the process is *Harrod-neutral*.

In Section V it was said that from 1953 to 1966 total recorded *R & D* rose from \$5.2 billions, the equivalent of 1.6 percent of private *GNP*, to \$22.2 billions, the equivalent of 3.3 percent. The industry-financed recorded *R & D* rose from \$2.2 billions, the equivalent of 0.7 percent of private *GNP*, to \$7.2 billions, the equivalent of 1.1 percent. How large is the “real” component of this increase?

A. One way of trying to answer this question is to go through the following moves:

a) We note that from 1953 to 1966 the number of *R & D* scientists and engineers, expressed in “full-time equivalents,” rose about 2.3-fold (from about 225,000 to about 520,000), and we now assume that the *total* real *R & D* input rose in the same proportion;

b) since during this period real private *GNP* (in 1958 prices) rose by 62 percent, we consider that the “real” component of the increase would have raised the ratio of total *R & D* cost to private *GNP* by 42 percent of the initial ratio (i.e., $230 \div 162 = 142$);

c) since in fact this ratio rose not by 42 percent but by 111 percent (from 1.57 to 3.31 percent), the relative-valuation component raised it by 49 percent of the initial ratio;

d) since for the *industry-financed R & D* the relative-valuation component was probably similar (49 percent), and since the ratio of the industry-financed *R & D* to private *GNP* actually rose by 60 percent (from 0.67 to 1.07 percent), the “real” component raised this ratio by merely 7 percent.

B. Another way of going about the matter would be to use Helen Milton's index *as if* it were an index of *R & D* input-prices—i.e., of cost of research per unit of *R & D* input in general—though in reality it is an index of cost of research per research worker (including supporting personnel). Use of this index requires a small amount of extrapolation, because the index stops in 1965. However with reasonable extrapolation this procedure leads to the conclusion that the relative-valuation component was somewhat larger than that estimated under A. The Milton index points to a *doubling* of cost-of-

research per worker from 1953 to 1966, while under *A*, we implied a 85 percent increase of the price of *R & D* input in general.⁶¹ With a 100 percent increase from 1953 to 1966, the "real" component would have raised the ratio of *total* recorded *R & D* to private *GNP* by 31 percent of the initial value of the ratio, and the relative-valuation component would have raised it by 61 percent.⁶² This 61 percent increase, brought about by the relative-valuation component, would have explained the *entire* increase of the ratio of *industry-financed* recorded *R & D* to private *GNP* (indeed, the relative-valuation component would have negligibly over-explained the 60 percent increase, see *A*, d)).

Both procedures—*A* and *B*—are crude, but it may be taken for granted that for the industry-financed *R & D* the "real" component was smaller than for total *R & D*. Indeed, for the industry-financed part the "real" component may have been negligible, though *both our procedures are suspect of somewhat underestimating the "real" component* because other *R & D* inputs probably rose in a higher proportion than *R & D* personnel.

The foregoing relates to the period 1953 to 1966. As for the pre-1953 decades, the records are very incomplete. Vanevar Bush's estimates suggest that the undeflated value ratio of "total" research costs to private *GNP* *trebled* from 1930 to 1944. From Helen Milton's cost-of-research index and the Department of Commerce's *GNP* deflator, we may infer that changes in relative values are responsible for only a small part of the increase of the proportion in question, and that even if we excluded the relative-value component we would still obtain a 2.5-fold increase of the ratio of research cost to private *GNP*. This is for the period 1930–44 which constitutes merely part of the span from 1929 to 1953. Judging by the data published

by the U.S. Bureau of Labor Statistics in its 1953 study of *Scientific Research and Development in American Industry*, the undeflated value-ratio seems to have increased by about 50 percent of itself from 1944 to 1952, with the relative-value component accounting for merely 13 percent increase and the "real" component for an increase of more than 30 percent.

2) *The Cost of Increased Per Capita Education.*

Turning to the per capita education inputs, one would assume that even without their growth in physical terms, they would become partly adjusted to the requirements of distributing *new* knowledge, but it is a fact that these inputs have been growing appreciably throughout the decades under consideration. Denison obtained a continuous and substantial increase over this period,⁶³ and a different way of getting at the problem leads to similar results.⁶⁴ We may put the average rate at which real per capita education inputs have been rising at roughly 2 percent; with a small upward adjustment for the value of education inputs in relation to other inputs we arrive at a rate of increase of between 2 percent and 3 percent. Teachers' salaries have recently been rising in relation

⁶³ Denison estimated that the average number of days of school attendance rose by 18.4 percent from 1920 to 1930; by 22.0 percent from 1930 to 1940; by 22.4 percent from 1940 to 1950; and by 20.0 percent from 1950 to 1960.

⁶⁴ If the rather narrowly defined *HEW* class of "education expenditures" is deflated by teacher's salaries, and the deflated expenditures are related to the size of the school-age population of both sexes (ages 5 to 21), the input so approximated is found to have risen by 25 percent per head of the relevant segment of the population from 1920 to 1930; by 6 percent from 1930 to 1940; by 33 percent from 1940 to 1950; by 14 percent from 1950 to 1960; and in a following period of merely five years' duration (1960 to 1965) by 10 percent. It so happens that this way of organizing the data leads to results not inconsistent with those of Denison from whose discussion one would also conclude—though on different grounds—that during these decades the average member of the labor force of any year had received about 2 percent more school education than the average member of the labor force a year earlier. In view of substantial differences in method, only rough correspondence can be expected with Denison's findings, but this is obtained.

⁶¹ An 85 percent increase would have raised the \$5.210 billions total *R&D* expenditure of the year 1953 to \$9.638 billions; and 2.3 times this amount (see *A*, a) equals \$22,167 billions. The 1966 *R&D* expenditure is estimated by *NSF* at \$22.220 billions.

⁶² This is instead of the 42 percent and 49 percent obtained for the two components, respectively, in *A*, b) and *A*, c).

to average incomes in the economy, though this is not true of all periods considered in this paper.

The question now arises as to the proportion of the private *GNP* to which this "2 percent to 3 percent increase" corresponds.

For 1953, *HEW* estimates its narrowly defined "educational expenditures" at an amount corresponding to 4 percent of the private *GNP*; for 1966 the same figure comes out at 6.8 percent. On the basis of Machlup's estimates for the second half of the 'fifties—estimates involving a few deductions from the *HEW* figures and a good many additions—the 4 percent and the 6.8 percent need to be multiplied by a factor of about 2.8 in order to arrive at education (knowledge-distribution) costs in the broad sense, the concept which is relevant from our point of view.⁵⁶ We arrive at 11 percent for 1953 and 19 percent for 1966. The text in Section V implies that "between 2 percent and 3 percent of 11 percent" may be regarded as 0.3 percent, and that "between 2 percent and

3 percent of 19 percent" may be regarded as 0.5 percent.

Using *R & D* statistics and education statistics additively is apt to lead to some amount of double-counting (the accounts are hardly kept with such consistency as to exclude this), but "imprecision" of the figure of 0.3 percent or of 0.5 percent resulting from this and from other reasons is very unlikely to affect the calculations presented in the paper.

APPENDIX C

(see Section V)

Listing the Unrecorded Costs of Technological Progress

1) Obsolescence of capital goods, as conventionally defined (i.e., including residential buildings).

2) Obsolescence of durable consumer goods as conventionally defined (i.e., excluding residential buildings). This would be a weighty item but it is a questionable candidate for inclusion because of the understatement of the quality-improvement of consumer goods in the productivity statistics.

3) Expenditures on *R & D* plant. These are not included in the "recorded *R & D*" figures of the paper, though incomplete numerical information is available on them. In some years the item may have corresponded to as much as 0.3 percent of private *GNP*.

4) "Technological" unemployment and underemployment. This is a doubtful item in the present context because it raises the question of the limitations of full-employment policies given alternative rates of technological progress. In some countries with a high rate of progress unemployment is exceedingly low.

5) The abandonment of methods of production which would remain associated with further learning by doing. This presumably is a small item because with given facilities *and for given models* learning by doing tapers off rapidly and the methods that are abandoned are usually the relatively old ones.

6) Unrecorded research workers with essentially the same characteristics as those

⁵⁶ See pp. 103 ff. in Machlup's book, the reference to which is found in Section V, *supra*; and cf also p. 83 in Machlup, *op. cit.*

Two remarks need to be added here. (1) The multiplier must have been different for different years but this can hardly have had an appreciable effect on the figure we are using in our yield-estimates for progress-generating inputs as a whole. (2) For our purpose some minor items should be omitted from those listed by Machlup on pp. 104-105 of *op. cit.*, but on the other hand, some roughly equivalent items should be added which Machlup interprets as knowledge expenditures but not as education costs and which he therefore does not list on pp. 104-105. For my purpose I omit from his education items the cost of tax-exemptions and the basic-training component of his military education-expenditures; and I replace these with the cost of books, periodicals and newspapers valued at prices charged to consumers (so valued because advertising expenditures do not belong in the field of our present concern).

Machlup, when relating education costs to *GNP* adds to measured *GNP* some items which he rightly regards as education costs but which are not included in the conventionally measured *GNP*. This I have not done, because for the present purpose, all costs should be related to the same output-base to which the measured productivity-increases are related.

The largest item included in Machlup's list but not in the *HEW* education expenditures consists of the earnings foregone by students and by the mothers of pre-school children.

recorded, but who for various reasons have not become included. Among these we have *inter alia* the "independent" inventors whose role in the innovating process remains important yet whose weight for our *cost calculus* is small.⁶⁶ Another group of considerable importance but of small weight for the present purpose consists of those engaged in market research, quality-control, and product-testing (these are not included in the *R & D* statistics).

The weight of some of the subcategories in item 6) has probably increased and that of others decreased over the past decades. The weight of those who have remained unrecorded merely because of the technical deficiencies or recording procedures must have somewhat decreased even if we compare the mid-fifties with the present.

7) Last but not least, the entrepreneurial input and its clerical complement, interpreted as the input required for preparing and making the top decisions on technological and organization change. This is obviously an 'elusive' item and one of small weight *in the present context*, though of crucial significance.

APPENDIX D

(see Section VI)

Reminder of Controversial Structural Problems in the Area of R & D-Policy

Somewhat more than 30 percent of the Federal *R & D* obligations for the fiscal year 1966 were those of *NASA*; about 45 percent consisted of those of the Defense Department (this undoubtedly included items not directly related to military objectives in the narrower sense); between 5 percent and 10 percent were obligations of the Atomic Energy Commission; and between 5 percent and 10 percent those of *HEW*, mainly for public health. Neither the main part of this paper nor the present Appendix contains detailed numerical information for the past three years because the preliminary data for these years are apt to become revised; the reader should therefore be reminded of the fact that

⁶⁶ See John Jewkes, David Sawers, and Richard Stillerman.

in 1966 space research reached a peak, though it has been a very weighty item of the government's *R & D* budget for the decade of the sixties as a whole. On my own list of priorities space exploration stands high and I value the Giant Leap of July 20, 1969 accordingly. But one should not pretend to express a professional judgment as an economist when attaching values to alternative amounts so allocated.

From 1953 to 1966 about two-thirds of the total recorded *R & D* was *financed* by the Federal government.⁶⁷ Industry *performed* about 70 percent of the total recorded *R & D*, while the remaining 30 percent was performed by the government itself and by other nonprofit institutions, including universities. The large discrepancy between the public-to-private ratio in *financing*, on the one hand, and in *performance* on the other, reflects the circumstances that the Federal government has typically financed about 55 percent of the *R & D* which was performed by industry (less in the first few years of the period 1953–1966, and slightly less also after 1966).

The typical share of Aircraft and Missiles in the total *R & D performed by industry* has been one-third or more, and the share of Electrical Equipment and Communication about one-fourth. This adds up to nearly 60 percent for the two leading performers among the industries. If we add Chemicals and Allied Products, Machinery (other than electrical), and Motor Vehicles—each of these three with shares falling somewhat short of 10 percent—we are up to about 85 percent of the total, for *five* "industries" as defined by the *NSF*.⁶⁸

⁶⁷ In the beginning of the period the government financed less than two-thirds, and after 1966 *somewhat* less.

⁶⁸ In a sense it is more revealing to add Petroleum Refining and Extraction and Professional and Scientific Instruments to the five leaders (which raises the joint share of the "leaders" so extended to somewhat more than 90 percent), because if we rank the industries by *R&D in relation to Value Added*, then these two industries are in the same range as the lower two of the top five. The *seven* leaders produce almost 17 percent of the U.S. national income.

We should note that the individual "industries" distinguished by the *NSF* are sometimes composites of

While, as was said, the typical government-financed proportion of the industry-performed *R & D* has been roughly 55 percent, the government-financed proportion of the *R & D* performed by Aircraft and Missiles has exceeded 80 percent (except for a recent small decline), and the government-financed proportion of the *R & D* performed by Electrical Equipment and Communication has amounted to about two-thirds, with the result that about 80 percent of the government funds used for the industrially performed *R & D* have had the purpose of financing the *R & D* activities of the *two* industries that lead the list of performers.

Concentration by industry is smaller for privately financed *R & D* than for the government-financed component, though the five leaders listed above remain the same industries for the privately financed component (not in the identical order and with a much more even distribution among them). The five account for about 75 percent of the privately financed performance. Concentration by industry has been noted frequently, and a look at the figures makes it self-explanatory why this degree of concentration should have been noted with critical overtones. On the other hand, we must remember that the government finances privately performed *R & D* *mainly* when the latter promotes the feasibility and the efficiency of activities to be performed in the public sector. The government could of course extend the scope of its research-financing, partly by a broader interpretation of this same objective, and partly by subsidizing the research of various industries on other grounds (e.g. because of favorable external effects expected of specific changes).⁵⁹ Indeed, "strong cases" of this sort may now be receiving more consideration than they have in the past. Rightly so, though any attempt to put such a scheme into effect *on a scale comparable to that of*

programs oriented to the needs of the government itself would carry political arbitrariness, and so-called favoritism, much further than is inevitable anyhow in an economy with a large public sector. One would hope that the possibility of policies involving less concentration by industry will be explored carefully, even if only limited success can be expected.

As for *R & D* concentration *within* individual industries—concentration of "research intensity" by firms—this too raises problems of concern to all of us believing in the merits of competition. The concern is legitimate, despite the fact that with respect to concentration *in this sense* the findings are less conclusive than with respect to concentration by industry. Over significant size-ranges, the relationship between size and research-intensity is very different in different industries, particularly in the region of firms of considerable size. Of the various methods of aggregation that may be tried, some lead to a presumption that "in the aggregate" there exist size-ranges in which research-intensity—*R & D* per unit of some measure of scale of operations—declines with rising size, and that research-intensity is pulled up again at the upper end of observed sizes—perhaps pulled up to a maximum—by a *very* small number of "largest" companies operating in specific industries.

In view of the magnitude of the *R & D* totals—\$22.2 billions for 1966 and probably about \$25 billions for 1968, i.e., of late about 3 percent of the *total GNP*—all these problems of composition and structure deserve considerable attention. We should develop our reaction to them in terms of our personal values. I will add that the largest part of the total *R & D* falls in the *NSF* category labelled "development"; the next largest, but distinctly smaller component is labelled "applied research"; and what is labelled "basic research" accounts for no more than about 15 percent (a proportion to which this component has gradually risen from less than 10 percent). I do not consider this classification particularly revealing, though it is essential to bear in mind that the prior insights from which "problem-oriented" research could fruitfully develop have all along resulted

"Census industries," and are not even always put together from Census industries defined on the same digit level.

⁵⁹ See Richard R. Nelson, Merton J. Peck, Edward J. Kalachek, Chapters 8 and 9.

from work that was not focused on "practical" objectives.

Considering that some observers minimize the significance of *R & D* for progress largely because of the alleged insignificance of the spillover effect originating in the present type of defense and space programs, I will end by noting that even in the peak year of expenditures on space-research, nearly one-half of the recorded *R & D* (a very large sum) fell in neither of these two categories.

REFERENCES

- K. J. Arrow, "Classificatory Notes on the Production and Transmission of Knowledge," *Amer. Econ. Rev., Proc.*, May 1969, 59, 29-36.
- G. S. Becker, *Human Capital*, Natl. Bur. Econ. Res., New York-London 1964.
- V. Bush, *Science, the Endless Frontier*, Washington 1944, p. 80.
- E. F. Denison, *The Sources of Economic Growth in the United States*, Committee for Economic Development, New York 1962.
- W. Fellner, "Measures of Technological Progress in the Light of Recent Growth Theories," *Amer. Econ. Rev.*, Dec. 1967, 5, 1073-99.
- , "Specific Interpretations of Learning By Doing," *J. Econ. Theory*, Aug. 1969, 1, 119-40.
- Z. Griliches and D. Jorgensen, "The Explanation of Productivity Change," *Rev. Econ. Stud.*, July 1967, 34, 249-83.
- J. Jewkes, D. Sawers, and R. Stillerman, *The Source of Invention*, New York-London 1958.
- J. W. Kendrick, *Productivity Trends in the United States*, Nat. Bur. Econ. Res., Princeton 1961.
- S. Kuznets, *Capital in the American Economy; Its Formation and Financing*, Nat. Bur. Econ. Res., Princeton 1961.
- F. Machlup, *Production and Distribution of Knowledge in the United States*, Princeton 1962.
- E. Mansfield, "Industrial Research and Development: Characteristics, Costs, and Diffusion of Results," *Amer. Econ. Rev. Proc.*, May 1969, 59, 65-72.
- H. S. Milton, "Cost-Of-Research Index 1920-1965," *Operations Research*, Nov.-Dec. 1966, 14, 977-91.
- J. Minasian, "Research and Development, Production Functions and Rates of Return," *Amer. Econ. Rev. Proc.*, May 1969, 59, 80-86.
- R. R. Nelson, M. J. Peck, and E. D. Kalachek, *Technology, Economic Growth and Public Policy*, The Brookings Institution, Washington 1967.
- R. C. Wasson, et al, "New Estimates of Fixed Business Capital in the United States, 1929-1965," *Surv. Curr. Bus.*, Dec. 66, 46.
- , "Fixed Business Capital in the United States, 1927-1968," *Surv. Curr. Bus.*, Feb. 1969, 49.
- Department of Health, Education and Welfare, *Health, Education and Welfare Trends*, Washington 1965.
- , *Projection of Educational Statistics to 1976-77*, Washington 1968.
- National Science Foundation, *American Science Manpower*, Washington
- , *Federal Funds for Research, Development and Other Scientific Activities*, Washington
- , *Review of Data on Scientific Personnel*, Washington
- , *Reviews of Data on Science Resources*, Washington
- , *Scientific and Technical Manpower and Resources*, Washington 1964.
- U.S. Bureau of Labor Statistics, *Scientific Research and Development in American Industry*, Washington 1963.

About the Growth Path of Firms

By ROMESH K. DIWAN*

The theory of the firm has been developed, by and large, in terms of static equilibrium. Even in this limited field, much of it rests on hypotheses that are rarely tested empirically. When one comes to dynamics, economists have been more preoccupied with problems of growth of an economy instead of the firm. Thus we have conjectures such as, the growth path of the economic system will pass through a turnpike, but we have hardly any conjectures as to how a small firm ends up as a big corporation. In this paper we offer a few conjectures dealing with the growth of a firm. Our inquiry is limited to only one facet of this growth process. The conjectures that we offer relate to the behavior of technological factors, namely: (1) elasticity of factor substitution, (2) technological impact on labor efficiency, and (3) biasedness of technological change.

The method we have employed is based on the celebrated and censured production function. However, we have made an attempt to compromise with the non-neo-classicists. The conjectures we offer are not the result of hypothetical formulations of the assumptions and the derivation of conclusions therefrom, but are instead implications of our empirical exercises. These are based on inductive rather than deductive method. This paper suggests that all three technological factors describe an inverted U-shaped curve in the technological factor and output (or firm size) space—the latter being the X -axis. In other words,

* Rensselaer Polytechnic Institute. This paper was presented at the December 1968 meetings of the Econometric Society at Evanston, Illinois under the title, "Firm Size, Labor Efficiency and Production Function." I thank the editor and an unknown referee for valuable suggestions.

these technological factors, contrary to the average costs, grow with the firm, reach a maximum and start falling off. These conjectures, particularly the one on the technological impact on labor efficiency, have relevance to the theory of growth recently enunciated by Nicholas Kaldor in 1967. They also suggest the reintroduction, in economic analysis, of concepts like "optimum scale" or Marshall's representative firm.

The plan of the paper is as follows. A general production function is outlined in Section I. Like Kaldor's technical progress function, a labor efficiency progress function is suggested in Section II and its implications with respect to technological characteristics are examined in Section III. Section IV formalizes the theory into a testable hypothesis. Sample size is discussed in Section V and the empirical results presented in Section VI.

I. The Production Function

We postulate a general homogeneous production function,¹ namely

$$(1) \quad x_0^{-\rho} = (A_1 x_1)^{-\rho} + (A_2 x_2)^{-\rho}$$

where x_i ($i=0, 1, 2$) refer to output, capital and labor, respectively. The inputs x_i ($i=1, 2$) are assumed to be measured in physical and quantity terms. The A_i ($i=1, 2$) are treated as variables instead of constants. One interpretation, there may be others, of these variables is that they define (or measure) capital and labor inputs, respectively, in efficiency (or quality) units as contrasted with x_i ($i=1, 2$) which

¹ This form has been used by Diwan (1968b), Ryuzo Sato and Martin Beckmann, and John Wise and Yong-Her Yeh.

measure the same inputs as physical (or quantity) units.

II. *Labor Efficiency Progress Function*

Since 1957, and earlier, when Robert Solow concluded that 90 percent of the increase in labor productivity in the non-farm sector of the U.S. economy in the period 1909-49 was due to the residual (technical change), a number of scholars have tried to account for similar estimates.² Edward Denison and Zvi Griliches used "more or less" accounting procedures to reduce the residual to a small figure by attaching values to the qualitative aspects of the inputs. Thus, Denison concluded that education accounted for as much as 20 percent.³ Murray Brown and Alfred Conrad, on the other hand, expressly used a Cobb-Douglas type production function and made labor productivity a function of education, research and development. By and large, these studies have been interested in an aggregative analysis of a sector or an economy. Their attention has been concentrated on the new variable in research and development expenditure. Edwin Mansfield has analyzed the effect of this new variable at the firm level.

Basically these scholars have been trying to take into consideration the idea that A_2 is a variable. Concentrating at the firm (not corporation) level, one has to ask the question about the determinants of labor productivity afresh.⁴ Common sense seemed to suggest that labor productivity

is higher in countries where capital input (per unit of labor) is higher. Wasn't Kaldor (1957) rationalizing this observation when he suggested the technical progress function?⁵ It will thus seem that labor productivity is a function both of capital intensity and education, research and development expenditure. However, education, research and development will effect the technical knowledge which is brought about by new investment which in its turn changes the capital intensity (of labor input). Thus, capital intensity will seem to be used in a *first approximation* approach. We, therefore, borrow Kaldors' technical progress function of vintage '57.⁶ Instead of accepting it in toto, we modify it according to our concept of the production function. It may be remembered that Kaldor does not use a production function and yet a linear approximation of his technical progress function is nothing but a linear Cobb-Douglas production function.

Kaldor argues that labor productivity increases with capital intensity of labor. In view of the other factors, we hypothesize that labor input measured in efficiency units is a function of the capital intensity of the labor input measured in physical units. Symbolically

$$(2) \quad \frac{DA_2}{A_2} = f\left(\frac{Dk}{k}\right)$$

where $k = (x_1/x_2)$ and D refers to the operation dt . This formulation is different from Kaldor's in so far as we use A_2 instead of his $Z = (x_0/x_2)$. Following Kaldor, we call it the "Labor Efficiency Progress Function."⁷ So far as the function f is concerned

² It may be mentioned that these estimates are heavily biased upward. Diwan (1968a) finds that if the production function is not linear, the residual is biased upwards on two counts.

³ Wilson and Lithwick followed the same analysis for Canada though they found that their conclusions were not necessarily similar.

⁴ Research in theory has been preoccupied with a model that treats technical change as an endogenous variable. As Solow (1967) rightly points out, it is still an unsettled question. In his "Learning by doing" Arrow (1962) has provided one hypothesis.

⁵ Nelson finds the explanation in the diffusion process of technical change.

⁶ The 1962 vintage of Kaldor and Mirrlees is more of a product differentiation in the context of a vintage model than a conceptual advance over the 1957 vintage.

⁷ It will be seen that by giving it a different name, we have differentiated our product. An important reason is that Kaldor may think differently, and it is not fair to use his term which has acquired specific meaning.

it will have the general properties suggested by Kaldor; such as $f' > 0$, $f'' < 0$, $f' \rightarrow 0$ as $Dk/k \rightarrow \infty$. Since our aim is empirical and we do not have the degrees of freedom available to a pure theorist, we take a simple case of the above function and define

$$(2') \quad \frac{DA_2}{A_2} = m \frac{Dk}{k}$$

fully realizing that in using (2') instead of (2) we lose much of the niceties of the function.

Before we go further let us find out what m is. From (2') we find that m is the elasticity of the labor input measured in efficiency units with respect to the capital intensity of the labor input in physical units. Since k is one among the many measures of technology, we will not be far wrong if we interpret m as a measure of the impact of technology on the efficiency of labor. From the general restrictions of the f function in (2), $m > 0$. The idea behind function (2) is that all capital, by and large, is complementary to the efficiency of labor. Expenditures on training the labor input, x_2 , on facilities provided to labor in terms of space, tools, machines, capacity to work, etc., has the effect of raising the efficiency of labor. On the other hand, labor measured in efficiency units is a substitute to the labor measured in physical units. They are substitutes in the sense that they compete for the same facilities. No wonder the oft quoted statement, "one efficient man can do the work of more than two persons."

Since m is an elasticity, it may be useful to find out the implications of different values of m ; particularly for $m > / < 1$. Now, $m = 1$ means that the growth rate of labor efficiency is equal to the growth rate of capital intensity (of the labor input). The meaning of $m > 1$ is that the labor efficiency grows faster than the capital intensity. This will imply that the new or

additional capital is complementary with labor efficiency. In other words, the new capital is helpful for a large productivity of labor. In somewhat general and loose terms we may interpret that technology is more or less productive according as $m \geq 1$. It may be added that we are only suggesting an interpretation of m and not asking the more important question; namely, what causes m to be greater or less than one?

Integrating over time and taking antilogs of (2') we get

$$(3) \quad A_2 = A_{20}k^m$$

where A_{20} is a constant of integration. We have thus expressed the technical progress function in terms directly useable in the production function.

III. Character of Technical Change

Substitution of labor efficiency progress function (3) into the production function (1) yields

$$(4) \quad x_0^{-\rho} = (A_1x_1)^{-\rho} + (A_{20}x_2)^{-\rho}(k)^{-m\rho}$$

This substitution has implications for the production function, for capital-labor substitution, and for biases in technical change. Let us take these one by one. Taking the production function implications first, let's assume for purposes of comparison that A_1 is a constant. It can be seen that in that case, equation (4) is exactly the production function derived by Marc Nerlove in his attempt to rationalize the empirical evidence of George Hildebrand and Ta-Chung Liu. Recently Yao-chi Lu and Lehman Fletcher have called it the Variable Elasticity of Substitution (VES), production function. As Lu and Fletcher have rightly pointed out, this function contains Constant Elasticity of Substitution (CES), Cobb-Douglas, and Leontief as special cases. Thus, $A_{20} = 0$ or $m = 1$ gives the Leontief Case; i.e. $x_0 = A_1x_1$, and $x_0 \rightarrow (A_1 + A_{20})x_1$,

respectively. The Cobb-Douglas case follows if $A_1=0$ so that $x_0=A'_{20} x_1^m x_2^{1-m}$. In case $m=0$, equation (4) dissolves into the CES, namely

$$x_0^{-\rho} = (A_1 x_1)^{-\rho} + (A_2 x_2)^{-\rho}$$

Regarding the implication in terms of the elasticity of factor substitution, let us define two measures of elasticity. It may be mentioned that the elasticity of factor substitution refers to the inputs measured in physical units not efficiency units. Conceptually, one could develop six ($\frac{6}{2}$) elasticities of factor substitution from the various combinations of the two factors measured in two different units; physical and efficiency. However, we restrict our analysis here to the conventional measures between the factors measured in physical units only. Let us call σ_n elasticity of substitution when technical change is neutral in Hicks' sense;⁸ σ_n may be defined as

$$(5) \quad \sigma_n = \frac{1}{1 + \rho}$$

It will be recognized that this is exactly the elasticity of substitution associated with the CES production function. It follows from the assumption $m=0$ and A_1 is a constant. On the other hand, the elasticity of substitution associated with equation (4), assuming A_1 is constant, may be called σ_b ; i.e. elasticity of substitution when technical change is of the biased type. σ_b is defined as

$$(6) \quad \sigma_b = \frac{1}{1 + \rho - \frac{m\rho}{S_1}}$$

where S_1 refers to the share of capital.

⁸ One may question if σ_n is a measure of the elasticity of substitution? The inclusion here does not mean that our answer to this question is yes. We have included it here because of its historical value, since in the literature, by and large, it has been treated as a measure of elasticity of substitution.

Equation (6) has been derived by Nerlove and Lu and Fletcher in their respective studies. It can easily be seen that σ_b is a variable while σ_n a constant. σ_b can be easily expressed as a function of S_1 as Ryuzo Sato and Martin Beckmann have done. σ_b can also be compared with Sato and Ronald Hoffman's cases of variable elasticity. Thus

$$\sigma_b = f(S_1);$$

$$\text{if } S_1 = g(t) \quad \text{we have } \sigma_b = f[g(t)]$$

$$\text{if } S_1 = h(k) \quad \text{we have } \sigma_b = f[h(k)]$$

One can develop other formulations as well.

From (5) and (6) we can find a relationship between σ_n and σ_b . This is given

$$(7) \quad \frac{m\rho}{S_1} = \left(\frac{1}{\sigma_n} - \frac{1}{\sigma_b} \right)$$

Since m and S_1 are positive (by assumption), the difference between these two measures will be determined by ρ . Thus we have

$$(8) \quad \text{if } \rho \begin{matrix} > \\ < \end{matrix} 0 \quad \text{then } \sigma_b \begin{matrix} > \\ < \end{matrix} \sigma_n$$

since

$$(9) \quad \rho \begin{matrix} > \\ < \end{matrix} 0 \quad \text{implies } \sigma_n \begin{matrix} < \\ > \end{matrix} 1$$

Condition (8) may be written in a different form; namely

$$(10) \quad \sigma_b \begin{matrix} > \\ < \end{matrix} \sigma_n \quad \text{when } \sigma_n \begin{matrix} < \\ > \end{matrix} 1$$

The interesting thing, about condition (10) is that the difference between σ_b and σ_n is of an equilibrating nature. Thus when $\sigma_n < 1$, σ_b is larger; thereby getting closer to 1, while if $\sigma_n > 1$, σ_b is smaller so that it is coming closer to 1. Also, we can analyze the bias in the measurement of the elasticity if it is estimated on the assumption that there is neutral technical change in

Hicks' sense when in fact there is biased technical change.

In terms of the implications for technical change, the substitution of the labor efficiency progress function (3) into the production function (1) introduces the possibility of biased technical change under the Hicksian criterion. Assuming factor inputs x_i ($i=1, 2$) are growing at the same rate, $[Dx_1/x_1 = Dx_2/x_2]$ technical change will be biased if $DF_1/F_1 \neq DF_2/F_2$, where DF_i and F_i ($i=1, 2$) refer to the changes in marginal productivity and the marginal productivity, respectively. From function (4) we have⁹

$$(11) \quad \frac{DF_1}{F_1} = (\rho + 1) \left(\frac{Dx_0}{x_0} - \frac{Dx_1}{x_1} \right) + \rho m \frac{S_2}{S_1} \frac{Dx_1}{x_1}$$

and

$$(12) \quad \frac{DF_2}{F_2} = (\rho + 1) \left(\frac{Dx_0}{x_0} - \frac{Dx_2}{x_2} \right) + \rho m \frac{Dx_2}{x_2}$$

Functions (11) and (12) are not exactly symmetrical. Let us define g as a measure of the biasedness of technical change,¹⁰

⁹ It may not be irrelevant to compare it with other production functions.

(1) For Cobb-Douglas

$$\frac{DF_i}{F_i} = \left(\frac{Dx_0}{x_0} - \frac{Dx_i}{x_i} \right) \quad (i = 1, 2)$$

(2) For CES (linear homogeneous)

$$\frac{DF_i}{F_i} = (\rho + 1) \left(\frac{Dx_0}{x_0} - \frac{Dx_i}{x_i} \right) \quad (i = 1, 2)$$

(3) For function (1)

$$\frac{DF_i}{F_i} = (\rho + 1) \left(\frac{Dx_0}{x_0} - \frac{Dx_i}{x_i} \right) - \rho \left(\frac{DA_i}{A_i} \right) \quad (i = 1, 2)$$

¹⁰ Our g is related to B of Ranis and Fei such that $g \geq 1$ according as $B \geq 0$; where B is defined as

$$B = \frac{DF_1}{F_1} - \frac{DF_2}{F_2}$$

such that $g \geq 1$ implies labor using (capital saving), neutral and labor saving (capital using) technical change, respectively.¹¹ From (11) and (12) we can evaluate g as

$$(13) \quad g = 1 + \frac{\rho m (S_1 - S_2)}{\rho S_2 - (\rho + 1) S_1}$$

where

$$(14, 15) \quad S_i = \frac{\partial x_0}{\partial x_i} \cdot \frac{x_i}{x_0}; \quad (i = 1, 2)$$

and

$$(16) \quad \sum_{i=1}^2 S_i = 1$$

It is seen that g is a function of factor shares, factor substitution and technological impact on labor efficiency. Very intuitively this is what it should be, as W. E. G. Salter rightly points out on page 135, "... it is impossible to conceive of technical advance without factor substitution."

In view of the Hicksian criterion, we find that in the particular case when $g=1$, the innovations are labor saving, neutral, and labor using, according as the growth rate of capital input is greater than, equal to, or less than the growth rate of the labor input; both in physical units. However, there are three cases when $g=1$. (i) When $m=0$. It will be noticed that this is the case of CES production function. In our theoretical scheme, it is the case when labor efficiency is not dependent upon capital; i.e. capital is not complementary to labor efficiency. (ii) When $\rho=0$. This is the classical case, since the production function (1) itself becomes degenerate. (iii) When capital and labor shares are equal. To demonstrate

$$(17) \quad S_1 = S_2 \text{ may be written as } F_1 x_1 = F_2 x_2$$

A simple manipulation gives

¹¹ It will be noticed that for the Cobb-Douglas and linear form of CES, $g=1$.

$$(18) \quad \frac{DF_1}{F_1} + \frac{Dx_1}{x_1} = \frac{DF_2}{F_2} + \frac{Dx_2}{x_2}$$

Thus

$$(19) \quad \frac{Dx_1}{x_1} > \frac{Dx_2}{x_2} \text{ implies } \frac{DF_2}{F_2} > \frac{DF_1}{F_1}$$

Thus, if the growth rate of capital input is greater than that of labor input, then the growth rate of marginal productivity of labor input has to be higher than that of capital input. This is possible only if the innovations are capital using or labor saving.

On the other hand, if $g > 1$, then even for neutral innovations, the growth rate of capital has to be higher than that of labor input. This is a rather unlikely case. In a society where capital is increasing faster than labor input, it would seem that a large part of it will be used for purposes of raising the efficiency of labor. Such a society will have $g < 1$.

An interesting result of the substitution of the function (3) in (1) is that the growth rate of output is not affected. This equation is

$$(20) \quad \frac{Dx_0}{x_0} = S_1 \frac{Dx_1}{x_1} + S_2 \frac{Dx_2}{x_2}$$

exactly the equation derived from Cobb-Douglas and linear homogeneous form of CES. The reason for this may be sought in the definition of the factor shares.¹²

IV. A Testable Hypothesis

Our theoretical formulation so far consists of only two relations; the production

¹² The growth equation from (1) is

$$\frac{Dx_0}{x_0} = S_1^* \left(\frac{DA_1}{A_1} + \frac{Dx_1}{x_1} \right) + S_2^* \left(\frac{DA_2}{A_2} + \frac{Dx_2}{x_2} \right)$$

The difference arises from the definition of S_i and S_i^* ($i=1, 2$). Thus

$$S_1 = S_1^* + mS_2^*$$

and

$$S_2 = (1 - m)S_2^*$$

function (1) and labor efficiency progress function (2). Together these relations contain two parameters, ρ and m . Both these parameters relate to the technological aspect of the production system. ρ is a parameter of capital-labor substitution and m a measure of the technological impact on labor efficiency; the inputs being measured in physical units. Both these relations are technological and do not include any economic considerations. To make a meaningful economic model, on the other hand, we need to introduce economic relationships. To do so, let us assume that the firm for which we have outlined these technological relations is in equilibrium. In that case, the marginal product of labor input (in physical units) is equal to its price, i.e.

$$(21) \quad F_2 = p_2$$

where p_2 refers to the price of labor input, relative to output. Equation (21), however, assumes a perfect labor market. This may be unrealistic. We can easily relax the condition above and assume instead

$$(22) \quad \frac{DF_2}{F_2} = \frac{Dp_2}{p_2}$$

so instead of (21) we derive from (22)

$$(23) \quad F_2 = cp_2$$

where c is a measure of market imperfections.

On substitution and simple manipulation equation (23) can be written in the equivalent form

$$(24) \quad z = b_0 p_2^{b_1} k^{b_2}$$

where $z = (x_0/x_2)$ i.e. labor productivity, and where

$$(25) \quad b_0 = \left(\frac{c}{1 - m} A_{20}^\rho \right)^{1/(1+\rho)}$$

$$(26) \quad b_1 = \frac{1}{1 + \rho} = \sigma_n$$

$$(27) \quad b_2 = \frac{mp}{1 + \rho}$$

so that

$$(28) \quad \rho = \frac{1}{b_1} - 1, \quad \text{and}$$

and

$$(29) \quad m = \frac{b_2}{1 - b_1}$$

It will be noticed that equation (24) is the form first used by Hildebrand and Liu. It has recently been used by Lu and Fletcher. As pointed out in Diwan (1968b), Yeong-Her Yeh has also used a similar relation. None of these studies, however, has provided a meaningful interpretation of the coefficients b_i ($i=0, 1, 2$). Yet equation (24) can be interpreted as an explanation of labor productivity (in physical units) in terms of the technological impact on labor efficiency and capital-labor substitution. In terms of our theory, the above hypothesis comes quite close to the explanation of increases in labor productivity suggested by Salter. To quote from page 143,

The analysis has suggested that the variation between industries in the extent of increases in labor productivity can be explained primarily by the uneven impact of three influences: (i) improvements in technical knowledge, (ii) potential economies of scale and the extent of their realization, and (iii) factor substitution. Although analytically distinct, these three influences are highly interrelated.

In our formulation we have not been able to separate out the effect of economies of scale. Salter also found these two to be inseparable.¹³ Our hunch is that the effect of economies of scale is absorbed in m .¹⁴

¹³ "These two sources of increased productivity are difficult to separate for newly discovered economies of scale are a very important aspect of improving technology" (p. 133).

¹⁴ It is interesting to note that Kaldor (1967) also takes a similar point of view in his theory of economic development. More specifically, "We cannot really separate the effect of economies of large scale . . . from

On the other hand, our analysis has the advantage of separating the effects of technical change and factor substitution.

Our aim is the formulation of conjectures about the growth path of the technological parameters in relation to the growth path of the firm. Therefore, we need to study the behaviour path of parameters ρ and m over the growth of the firm size. Having formulated a theory that explains productivity of the labor input in terms of technological factors, equation (24), one possible method is to postulate assumptions which allow us to determine the growth path of these parameters as a function of the growth path of the firm. The second method is to estimate the possible values of these parameters for large, small, and medium sized firms and formulate our conjectures from this evidence. We propose to follow the second method. To do so, we need to establish a priori expectations about the statistical values of the parameters. From our theory, we know that (i) $0 < m < \infty$, (ii) $-1 < \rho < \infty$, (iii) $g > 0$, (iv) $\rho \neq 0$. From these conditions, we can lay down a priori expectations on the parameters b_i ($i=1, 2$) in equation (24). Thus we expect that (i) $b_1 \neq 1$, (ii) $\infty > b_1 > 0$ and (iii) $1 > b_2 > \infty$. Unless our statistical estimates fall within range and satisfy these conditions, our results will not be acceptable for interpretation by our theory. It may be added that the condition $b_1 \neq 1$ is a rather stringent one.

V. Sample and Subsample

For the purpose of estimation one needs data at the firm level and for different firm sizes. We have utilized the data collected by McGraw-Hill and described by Robert Eisner and Diwan (1965). The only difference between the data we have used and those used by Eisner, is that ours is a smaller subset. Our data refer to 500 manu-

those which are due to irreversible improvements of technology associated with a process of expansion" (p. 13).

facturing firms for the period 1955, 1956, and 1957, selected on a random basis. In view of the fact that some information was not available for some firms, we were left with the following sample:

TABLE 1—THE SAMPLE

Year	Number of Firms
1955	276
1956	346
1957	273

An attempt was made to divide the sample into large, medium, and small firms. The firms were arranged in a descending order of "sales." Mansfield and Frederick Scherer have also used sales as an index of size.

There was no marked dividing line, except for 1956 and some parts of 1957, between the firms so that it was difficult to set up a simple criterion for distinction between large, medium, and small firms. The firms presented, more or less, a continuum. Since we had the restriction that no subsample should contain only a few firms, we ended by breaking up the sample into subsamples on a more or less arbitrary basis. For 1956 alone, our efforts were successful. The distribution of the firms into the subsamples is given in Table 2.

TABLE 2—DISTRIBUTION INTO SUBSAMPLES

Year	Number of observations in subsample					
	1	2	3	4	5	Total
1955	100	100	76	—	—	276
1956	25	100	100	96	25	346
1957	25	98	79	71	—	273

Since the firms have been arranged on a descending order of sales, the first subsample contains the largest firms, and the last the smallest. Actually, for 1955 the firms are not much different in the subsample 2 and 3.

VI. Firm Size and Structural Parameters

The test of the theory lies in the estimation of relation (24). Expressing capital letters for logarithms, (24) may be written as:

$$(30) \quad Z = b_0 + b_1 P_2 + b_2 K + U$$

U is a random term. For purposes of estimation by least squares it is assumed that U has the necessary properties.¹⁵

Relation (29) has been estimated for all the subsamples and the results are presented in Table 3. It will be seen that the estimates of b_1 and b_2 are highly significant

¹⁵ To analyze U , however, it is necessary to study the composition of the variables Z , P_2 and K . In view of the nature of our data, these variables are likely to have errors of measurement. Let us denote the theoretical variables by (*) and assume that the observed variables are related linearly as below:

$$(1),(2) \quad X_i^* = X_i + v_i \quad (i = 0, 1)$$

$$(3) \quad P_2^* = P_2 + v_2$$

It will be noticed that we assume that $X_2^* = X_2$, since the number of workers are liable to be recorded reasonably accurately. v_i are errors of measurement. Assuming (24) holds for starred values, replacing these by non-starred (observed) ones, we find that:

$$(4) \quad \log u = \log u_0 - b_1 \log u_2 - b_2 \log u_1$$

where

$$(5) \quad u_0 = \left(1 + \frac{v_0}{X_0}\right)$$

$$(6) \quad u_1 = \left(1 + \frac{v_1}{X_1}\right)$$

$$(7) \quad u_2 = \left(1 + \frac{v_2}{P_2}\right)$$

To assume that U has the usual properties of zero mean and constant standard deviation implies a number of assumptions. These are:

$$(8) \quad E(U_0) = E(U_1) = E(U_2) = 0$$

$$(9) \quad \text{Var } U_0 = k_0; \quad \text{Var } U_1 = k_1; \quad \text{Var } U_2 = k_2$$

$$(10) \quad \text{Cov } (U_0 U_1) = \text{Cov } (U_1 U_2) = \text{Cov } (U_0 U_2) = 0$$

$$\text{Cov } (P_2 U_0) = \text{Cov } (P_2 U_1) = \text{Cov } (P_2 U_2) = 0$$

$$\text{Cov } (Y U_0) = \text{Cov } (Y U_1) = \text{Cov } (Y U_2) = 0$$

It may be added that assumptions in the system (10) are rather stringent. However, as a result of (8) and (9) we have $E(U) = 0$ and $\text{Var } U = \Sigma k_i$, ($i = 0, 1, 2$). Also we get relation (29) in terms of observed (nonstarred) values instead of (24) which is expressed in theoretical (starred) variables.

TABLE 3: $Z = b_0 + b_1P_1 + b_2K + U$

Year and Firm Size	Observations	b_0	b_1	b_2	R^2
1955					
1	100	0.3108 (.1553)	0.8589 (.0215)	0.1253 (.0161)	.98
2	100	-.2139 (.1864)	0.9245 (.0260)	0.1131 (.0234)	.98
3	76	-.0003 (.2551)	0.8120 (.0249)	0.2045 (.0264)	.96
1956					
1	25	0.7774 (.2564)	0.7449 (.0431)	0.1893 (.0272)	.99
2	100	0.2345 (.1696)	0.8665 (.0218)	0.1242 (.0176)	.97
3	100	-.1546 (.1711)	0.9318 (.0263)	0.0997 (.0219)	.97
4	96	-.4636 (.2074)	0.9293 (.0272)	0.1360 (.0204)	.96
5	25	1.0331 (.8748)	0.6399 (.0545)	0.2638 (.0957)	.89
1957					
1	25	0.9145 (.2455)	0.6407 (.0448)	0.1589 (.0332)	.98
2	98	0.3097 (.1607)	0.8894 (.0189)	0.0931 (.0151)	.97
3	79	-.0006 (.1845)	0.9491 (.0233)	0.0651 (.0237)	.98
4	71	1.2059 (.3758)	0.6329 (.0322)	0.2370 (.0393)	.90

(Figures in parentheses refer to standard errors.)

z is the output-labor ratio. Output is measured as sales plus change in inventories. Labor is measured by the number of employees of the firm. k is the capital-labor ratio and capital is measured by the gross value of plant and equipment adjusted by the actual utilization rate. p_1 is measured by the relationship $(1 - S_1)z$. Note that $Z = \log z$, etc.

for all regressions. The important result is that b_1 is significantly different both from zero and one. Also, the estimates of b_i ($i=1, 2$) are significantly different for the different firm sizes. The coefficient of multiple regression is quite high. Our theory, therefore, provides a meaningful explanation of the variations in the physical labor productivity over different firms. We may

also compare these results with those of Hildebrand and Liu, and Lu and Fletcher. Hildebrand and Liu find that coefficients b_1 and b_2 are rarely significant together. Thus, from their results for all employees they get five regressions out of seventeen where both b_1 and b_2 are significant. When they use production workers only, this number falls down to two. Lu and Fletcher

TABLE 4—ESTIMATES OF THE PARAMETERS

Year and Firm Size		Parameters			
		σ_n^a	σ_b^b	m^c	g^d
1955					
	1	.8589	.9945	.8885	.8850
	2	.9245	1.0547	1.4969	.8960
	3	.8120	1.0446	1.0887	.8105
1956					
	1	.7449	.9409	.7421	.8268
	2	.8665	1.0025	.9078	.8892
	3	.9318	1.0456	1.4619	.9088
	4	.9295	1.0918	1.9224	.8749
	5	.6399	.8936	.7396	.7492
1957					
	1	.6407	.7760	.4422	.8543
	2	.8894	.9901	.8423	.9132
	3	.9491	1.0216	.9798	.9408
	4	.6529	.8795	.6828	.7793

^a Defined in equation (5).

^b Defined in equation (6).

^c Defined in equation (21).

^d Defined in equation (13).

did not present the results, although they maintain that b_2 is significantly different from zero in a number of cases. Incidentally, both these studies were preoccupied with testing of b_2 significantly different from zero; since this provides a test if the production function is of the *CES* or *VES* variety. None of these studies test if b_1 is significantly different from one.

In order to estimate the basic parameters σ_n , σ_b , m and g we need values for S_1 .¹⁶ In view of the complexity of the relationship between the basic parameters and b_i ($i=0, 1, 2$), it was not possible to estimate the standard errors. A cursory and a superficial glance may give one a wrong impression, yet if one looks closely, it is not so serious a deficiency as to vitiate the conclusions. First, the standard errors

of the b_i ($i=1, 2$) are of "more or less" the same magnitude. Second, since standard errors of basic parameters are functions of the standard errors of the b_i ($i=1, 2$), it is improbable that these standard errors will effect the pattern over firm size; a pattern generated by estimates of b_i ($i=1, 2$).

The estimates for σ_n , σ_b , m and g are presented in Table 4. Taking individual parameters first: the estimates of σ_b are consistently higher than those of σ_n . Actually this follows from the fact that $b_2 > 0$. Nerlove has estimated σ_b for the results in Hildebrand and Liu. He also finds that $\sigma_b > \sigma_n$. Lu and Fletcher do not give any estimates of σ_b .¹⁷ All they mention is that these differ from σ_n . So far as the estimates of m and g are concerned, I am afraid, there

¹⁶ These estimates have been estimated by the method now known as the Klein method; they seem to have an upward bias. These are consistent with the estimates from other methods; simplistic and complex. Details are provided in Diwan (1965).

¹⁷ It is a pity Sato and Hoffman did not record the values of variable elasticity of substitution. One wishes the editors would have asked for this information since it would not have implied any extra work or space. Who can question the wisdom of editors?

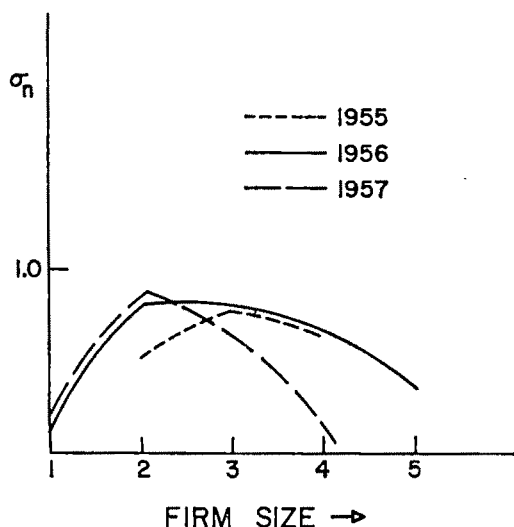


FIGURE 1

are no precedents. Since m is an elasticity, one's a priori expectation is that it will not be too far from one. There is no reason for such an expectation except that majority of the quantitative estimate for various elasticities (like elasticity of demand, elasticity of substitution, elasticity of supply, etc.) stay around 1. Judged from such expectations, the values of m in Table 4 seem to be of the right magnitude. So far as g is concerned, we do have a priori expectation, namely, technical change in U.S. manufacturing is labor saving. From this expectation $g < 1$ and the values of g in Table 4, therefore, seem quite reasonable. Another way to have a feel for these magnitudes, is to compare them with the results of Hildebrand and Liu. I have estimated these parameters for Hildebrand and Liu¹⁸ and lo and behold, the magnitudes are similar. g is less than one and between .87 and .98 for food and kindred, furniture and fixtures, paper and pulp, chemicals, leather and leather products, and stone, clay, and glass. For all others it is greater than one, though not greater than 1.12. One must, however, remember

that standard errors in Hildebrand and Liu are rather large.

VII. Conclusions

Our theory explains reasonably well the variations in the physical labor productivity. The estimates of b_1 ($i=1, 2$) are significant and of the same magnitude as the estimates from other studies. Estimates of our basic parameters look reasonable. We may now come to the central question as to the way our basic parameters move with the growth of the firm from a small to a large one. In other words, what is the path of these parameters associated with the growth of the firm. To answer, we need to look at these parameters for individual years and over different firm sizes. We find that all these parameters have the same growth path, though their magnitudes vary. We have described these in the following Figures 1-4. What we find is that for large and small firms (first and last subsamples) the estimates are low, and for medium firms these are high. Thus, if we measure on the X-axis the size of the firm, either in ascending or descending order, and on the Y-axis the value of the parameter, then all these parameters describe an

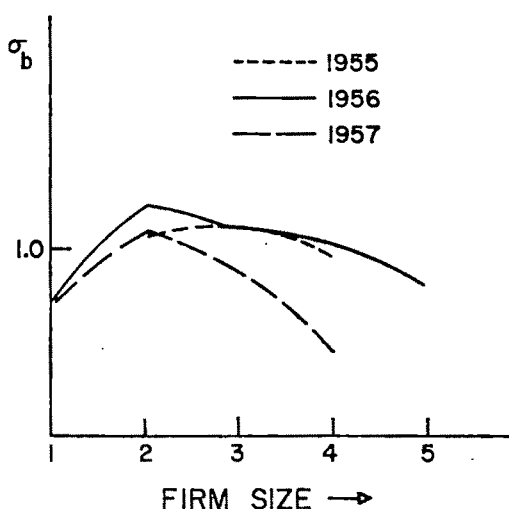


FIGURE 2

¹⁸ These results are available on request.

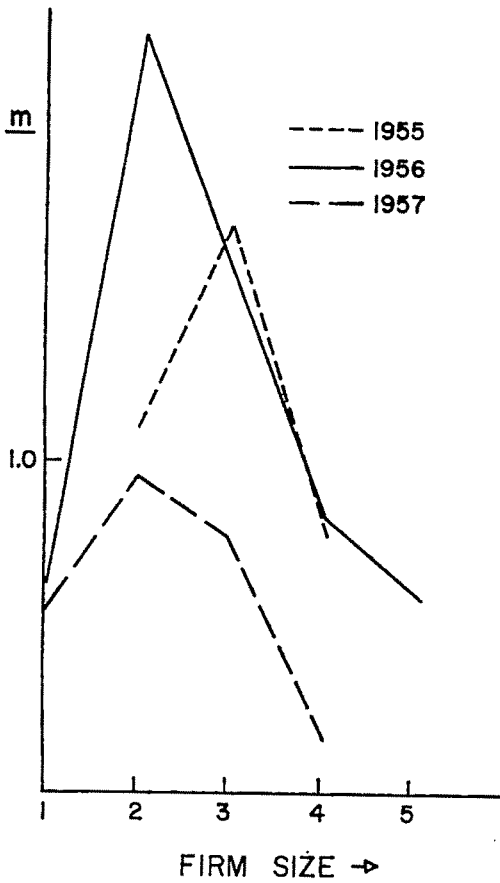


FIGURE 3

inverted U-shaped curve which first rises then falls, reaching a maximum somewhere in between. This pattern seems to be stable and does not change if we use estimates of S_1 by another method.¹⁹ One may describe

¹⁹ As an alternative, we have estimated S_1 by the following method: Assume the Cobb-Douglas:

$$(1) \quad x_0 = ax_1^{d_1} x_2^{d_2}$$

and

$$(2) \quad (d_1/d_2) = h = \text{antilog} \frac{1}{n} \sum \log \frac{P_1 x_1}{P_2 x_2}$$

so that

$$X_0 = A + d_2(X_2 + hX_1) + U$$

and impose the restriction

$$(3) \quad d_1 + d_2 = 1, \text{ so that } S_1 = d_1$$

The following Table gives the estimates for σ_b and g

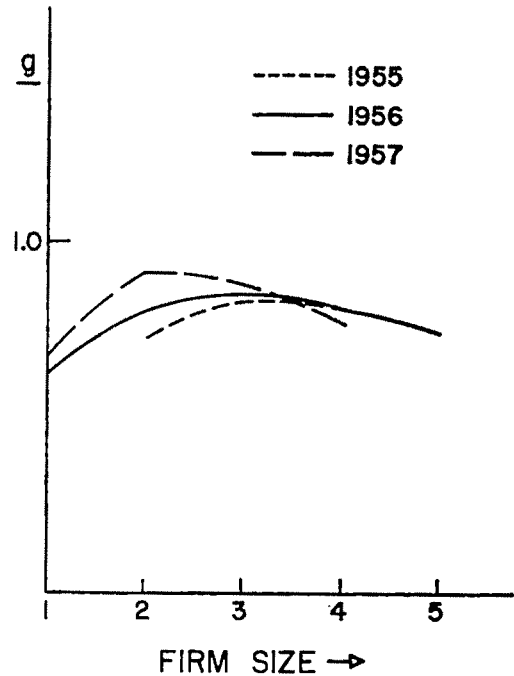


FIGURE 4

the process as follows. As the firm grows from small to medium, all these technological factors are becoming more effective. There is greater factor substitution thus allowing greater flexibility and maneuverability. The impact of technical change is to make labor more efficient so that tech-
using S_1 as derived from (3):

TABLE 2n. ESTIMATES OF σ_b AND g

Year and Firm Size	σ_b	g
1955		
1	1.0037	.8923
2	1.0591	.8998
3	1.0532	.8157
1956		
1	.9483	.8320
2	1.0056	.8916
3	1.0485	.9047
4	1.1011	.8812
5	.9414	.7786
1957		
1	.7761	.8543
2	.9903	.9152
3	1.0228	.9418
4	.8985	.7892

nology is more productive. One has a hunch that there would be economies of scale compounded in this process. The technical change is becoming less biased implying a stable or efficient matching of resources. Yet if the firm still grows in size, these advantages are lost. On a more general level, these results agree with those of Mansfield²⁰ and Scherer.

REFERENCES

- K. J. Arrow, "The Economic Implications of Learning by Doing," *Rev. Econ. Stud.*, June 1962, 29, 155-73.
- , H. Chenery, B. Minhas, and R. Solow, "Capital-Labor Substitution and Economic Efficiency," *Rev. Econ. Statist.*, Aug. 1961, 43, 225-50.
- M. Brown and A. H. Conrad, "The Influence of Research and Education on C.E.S. Production Relations," in M. Brown, ed., *The Theory and Empirical Analysis of Production*, Nat. Bur. Econ. Res. Stud. in *Income and Wealth*, Vol. 31, New York 1967.
- E. F. Denison, *The Sources of Economic Growth in the U.S. and the Alternatives Before Us*, Washington 1962.
- R. K. Diwan, "Specification and Estimation of Production Functions for U.S. Manufacturing," unpublished doctoral dissertation, Univ. Birmingham, 1965.
- , 1968a, "A Bias in the Measurement of Technical Change," *J. Finance Quant. Anal.*, 1968, 3, 471-77.
- , 1968b, "Returns to Scale in Indian Industry: A Comment," *The Econometric Annual of Indian Econ. J.*, 1968, 496-500.
- R. Eisner, "Capital and Labor in Production: Some Direct Estimates," in M. Brown, ed., *The Theory and Empirical Analysis of Production*, Nat. Bur. Econ. Res. Stud. in *Income and Wealth*, Vol. 31, New York 1967.
- Z. Griliches, "Production Functions in Manufacturing: Some Preliminary Results," in M. Brown, ed., *The Theory and Empirical Analysis of Production*, Nat. Bur. Econ. Res. Stud. in *Income and Wealth*, Vol. 31, New York 1967.
- G. H. Hildebrand and T. C. Liu, *Manufacturing Production Functions in the U.S. 1957*, Ithaca 1965.
- N. Kaldor, "A Model of Economic Growth," *Econ. J.*, Dec. 1957, 67, 591-624.
- , *Strategic Factors in Economic Development*, Ithaca 1967.
- and J. A. Mirrlees, "A New Model of Economic Growth," *Rev. Econ. Stud.*, June 1962, 29, 174-92.
- Y. Lu and L. B. Fletcher, "A Generalization of the C.E.S. Production Function," *Rev. Econ. Statist.*, Nov. 1968, 50, 449-52.
- E. Mansfield, *Industrial Research and Technological Innovation: An Economic Approach*, New York 1968.
- R. R. Nelson, "A Diffusion Model of International Productivity Differences in Manufacturing Industry," *Amer. Econ. Rev.*, Dec. 1968, 58, 1219-48.
- M. Nerlove, "Recent Empirical Studies of the C.E.S. and Related Production Functions," in M. Brown, ed., *The Theory and Empirical Analysis of Production*, Nat. Bur. Econ. Res. Stud. in *Income and Wealth*, Vol. 31, New York 1967, pp. 25-50.
- G. Ranis and J. Fei, "Innovational Intensity and Factor Bias in the Theory of Growth," *Int. Econ. Rev.*, May 1965, 6, 182-98.
- W. E. G. Salter, *Productivity and Technical Change*, New York 1966.
- R. Sato and M. J. Beckmann, "Neutral Inventions and Production Functions," *Rev. Econ. Stud.*, Jan. 1968, 35, 57-65.
- R. Sato and R. F. Hoffman, "Production Functions With Variable Elasticity of Factor Substitution: Some Analysis and Testing," *Rev. Econ. Statist.*, Nov. 1968, 50, 453-60.
- F. M. Scherer, "Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions," *Amer. Econ. Rev.*, Dec. 1965, 55, 1097-1125.
- R. M. Solow, "Technical Change and the Aggregate Production Function," *Rev. Econ. Statist.*, Aug. 1957, 39, 312-20.
- , "Some Recent Developments in the

²⁰ Mansfield finds that in steel the inventive output follows a similar growth path.

- Theory of Production," in M. Brown, ed., *The Theory and Empirical Analysis of Production*, Nat. Bur. Econ. Res. Stud. in *Income and Wealth*, Vol. 31, New York 1967.
- T. A. Wilson and N. H. Lithwick, *The Sources of Economic Growth*, Studies of Royal Commission on Taxation, No. 24, Ottawa 1968.
- J. Wise and Y. H. Yeh, "Econometric Techniques of Analyzing Wage and Productivity Differentials with Applications to Manufacturing Industries in U.S.A., India and Japan," Paper presented at the 1965 Econometric Society Meeting at New York.
- Y. H. Yeh, "Returns to Scale in Indian Industry," *Econometric Annual of Indian Econ. J.*, 1966, 275-291.

Portfolio Choice, Investment, and Growth

By DUNCAN K. FOLEY AND MIGUEL SIDRAUSKI*

One of the basic postulates of monetary theory and policy is that an increase in the money supply raises income by exerting at least temporary downward pressure on the rate of interest. Since it is generally believed that the savings rate is rather unresponsive to changes in the rate of interest, the key to this postulate, and to the belief that monetary policy can affect aggregate demand and its composition, seems to be the existence of a negative relationship between the demand for investment goods and the market rate of interest.

The micro-foundations of such a relationship, however, are very weak. It is well-known from the work on investment theory by Trygve Havvelmo and others that an analysis of profit maximizing behavior on the part of firms can at most offer a relationship between the desired stock of capital and the rate of interest, but not a relationship between the desired rate of increase in the stock of capital

and the rate of interest. The demand for investment cannot be derived from the demand for capital. Demand for a finite addition to the stock of capital can lead to any rate of investment.

Recent attempts to derive a marginal efficiency of investment schedule from a firm's profit maximizing behavior have relied largely on the assumption that there are costs to adjusting the actual to the desired stock of capital at a fast rate; that is, the addition to the capital stock which results from each additional dollar spent on investment diminishes with the level of investment.¹ Alternatively, they have simply assumed that once a firm finds out what its optimum stock of capital is (if it exists), it then adjusts the actual to the desired stock with a lag (see Dale Jorgenson); the simplest of these models being the one in which the rate of investment is a linear function of the difference between the desired and the actual stock of capital. In some models of economic growth, this theory requires a permanent gap between the desired and the actual stocks of capital in a state of steady growth to generate the investment necessary to maintain a constant capital-labor ratio over time when population is growing (see Richard Nelson).

It is our purpose in this paper to point out another mechanism which can determine the aggregate rate of investment. This mechanism depends on socially rising costs to investment rather than rising costs to individual firms. We present a very simple model of growth in which the level of investment is determined jointly

* Duncan Foley is assistant professor of economics at the Massachusetts Institute of Technology. Miguel Sidrauski died on September 1, 1968. A preliminary version of this paper was presented at the Conference on Money and Economic Growth held at Brown University, June 1968. We would like to acknowledge the substantial contribution of Stanley Fischer to this paper. Mr. Fischer participated in many of our preliminary conversations about its substance and shape, and undertook extensive editorial work in the later stages. We would also like to thank James Mirrlees, Don Patinkin and Carl von Weizsäcker for our very helpful discussions with them on the subject of this paper. The line of thought developed in this paper draws heavily on the published and unpublished work of James Tobin, especially in the treatment of asset demands and prices.

Duncan Foley's work was supported by the National Science Foundation (grant number GS 1585). The views expressed here are the authors' sole responsibility, and do not reflect those of the National Science Foundation, the department of economics, nor those of the Massachusetts Institute of Technology.

¹ For this approach see R. Eisner and R. H. Strotz, J. P. Gould, and Arthur Treadway.

by the interaction of the stock of capital, its demand, which is treated as an integral part of the demand for assets by wealth owners, and the flow supply of investment goods as determined by the producers of capital goods. Even if the price of capital is such that the marginal return to the existing stock of capital makes the stock instantaneously optimal, this does not necessarily imply that this price makes it optimal for producers of capital goods not to produce any capital goods at all.

One advantage of this approach is that it permits firms and consumers to be in equilibrium at each instant and still generates new investment. The relative importance of rising firm costs and rising social costs in determining rates of investment in various sectors is, of course, a question which empirical studies must answer.

We believe that our model of investment corresponds quite closely to Keynes' vision of the investment process. In his response to comments on the *General Theory* he states:

The owner of wealth, who has been induced not to hold his wealth in the shape of hoarded money, still has two alternatives between which to choose. He can lend his money at the current rate of money-interest or he can purchase some kind of capital-asset. Clearly in equilibrium these two alternatives must offer an equal advantage to the marginal investor in each of them. This is brought about by shifts in the money-prices of capital-assets relative to the prices of money-loans. The prices of capital-assets move until, having regard to their prospective yields and account being taken of all those elements of doubt and uncertainty interested and disinterested advice, fashion, convention, and what else you will, which affect the mind of the investor, they offer an equal, apparent advantage to the marginal investor who is wavering between one kind of investment and another. . . .

Capital-assets are capable, in general, of being newly produced. The scale on

which they are produced depends, of course, on the relation between their costs of production and the prices which they are expected to realize in the market. Thus if the level of the rate of interest taken in conjunction with opinions about their prospective yield raise the prices of capital-assets, the volume of current investment (meaning by this the value of the output of newly-produced capital-assets) will be increased; while if, on the other hand, these influences reduce the prices of capital-assets, the volume of current investment will be diminished. [p. 188]

Given the state of technology and the stock of capital which is a result of past saving, the only variable that is unequivocally related to the level of investment is the price of capital goods. It is this emphasis, together with the notion of instantaneous stock equilibrium in the asset markets determining asset prices and yields, to be discussed below, which we share with Keynes' description quoted above. When the price of capital goods is high, the level of investment is also high; while if the price of capital is low, investment is also low. This would seem surprising if we look at the investment function as an *ex ante* demand relationship because one would then think that the higher is the price of capital, the lower will be the level of investment. This result is not at all surprising if we think of the investment function as an *ex post* supply relationship. The higher is the relative price of capital, the larger will be the share of the economy's total resources that producers will find profitable to allocate to the production of investment goods, and therefore, the higher will be the rate of capital accumulation.

We do not, however, assume the existence of a negative relationship between the desired level of investment and the market rate of interest and we show that such a relationship may not even exist between the actual, *ex post*, rate of invest-

ment and the market rate of interest.² Other things equal, an increase in the rate of interest lowers the demand for capital on the part of wealth owners, decreasing on this account the equilibrium price of capital and thus lowering the optimum level of output in the investment goods sector. Other things, however, are not usually equal. The rate of interest is itself an endogenous variable of the system. A change in the rate of interest implies that some other variable in the system has changed, and the change in this variable may have increased the demand for capital.

Although we conclude that there may be neither an *ex ante* nor an *ex post* negative relationship between the level of investment and the market rate of interest, this does *not* imply that the effects of monetary policy on aggregate demand or its composition are unpredictable. On the contrary, we shall prove that monetary policy is able to change aggregate demand and its composition by affecting the price of capital at which wealth owners are willing to hold the existing stock of capital. Monetary and fiscal policy jointly not only are able to affect the rate of growth by changing the relative price of capital, but they can also succeed in maintaining a constant price level while at the same time determining the economy's rate of growth.

Once the equilibrium level of investment is determined, we go on to consider the process by which capital and other assets find room in private portfolios. Instantaneous market equilibrium ensures that, given the state of expectations, wealth owners are willing to hold the existing stocks of assets and that their desired and actual savings are equal. It does not, however, guarantee that at the market clearing prices the value of the addition to each of

their assets is equal to the value of the desired additions. That is, our equilibrium theory of investment does not assure the equality of the flow supplies and demands for each of the assets at the current equilibrium prices. Since markets clear at each moment of time, the process of wealth accumulation will generally require prices to change over time. Actual price changes and expected price changes affect demand and supply schedules which in turn modify equilibrium prices over time.

In the last section of this paper we study the full dynamic behavior of the economy over time using a particular expectations hypothesis which has been employed previously by other authors. In this model we show how actual and expected price changes interact while portfolios are adjusted to make room for the newly supplied assets. In the context of this model, we also study how monetary and fiscal policy are able to stabilize aggregate demand so as to maintain a constant price level and how changes in the mix of fiscal and monetary policy affect the economy's long-run capital stock by modifying the price of capital at which wealth owners are willing to hold the existing stock of capital at any point in time.

I. *The Demand for Capital Services*

Language itself presents several traps in thinking about investment, especially confusion between the demand for capital, the demand for capital services, and the demand for investment. It seems to be appropriate to sort out these concepts at the start.

The simplest notion is the demand for capital services. This is analogous to the demand for the services of any factor of production, and arises indirectly from the demand for output. Capital services have the dimension of a rate or flow: the use of such and such a machine or building for

² J. G. Witte takes a similar approach to ours but concludes with an *ex post* relationship between the rate of investment and the market rate of interest.

one hour, week, or month. The price which is determined in the market for capital services will be called the rental rate. It is measured in value units per hour, month, or year, like the wage rate.

Producers will demand at any wage-rental rate combination those amounts of capital and labor services which will equate their marginal products of capital and labor to those rates. In thinking about equilibrium in the market for capital services, we are free to ignore all matters concerned with the ownership of capital. A producer trying to equate marginal products to the prices of factor services is making a decision which is entirely divorced conceptually from consideration of owning capital. It may be that taxes and other market imperfections distort producer choices but in principle the producer in deciding his production plan does not care who owns the capital. He decides entirely on the basis of the market rental rate, which will be the same whether he, his competitor, or the King of Siam owns the machines and buildings.

A simple model which exemplifies these distinctions is the well-known two-sector model of production. One sector produces investment goods, I , and the other produces consumption goods, C . We take the consumption good to be the numeraire and denote the consumption price of capital goods by p_k (see T. M. Rybczynski).

In this model the demand for the services of capital and labor will be determined by the familiar marginal productivity conditions, which bear no relationship to the ownership of factors. The rentals to capital in the two sectors under competitive conditions are the marginal value products of capital:

$$(1.1) \quad r_C = f'_C(k_C) = r$$

$$(1.2) \quad r_I = p_k f'_I(k_I)$$

If, as assumed, capital is perfectly mo-

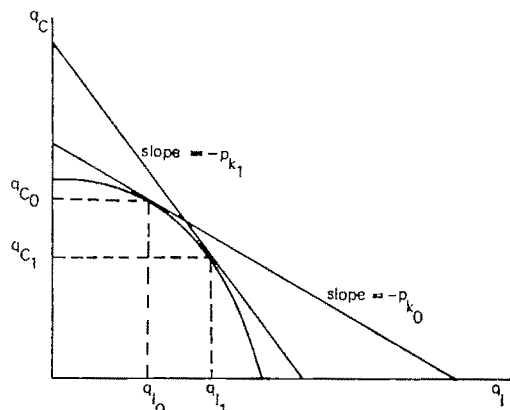


FIGURE 1

bile between the two sectors, these two must be equal at each moment of time:³

$$(1.3) \quad f'_C(k_C) = p_k f'_I(k_I).$$

In the two-sector model, the per capita supplies of consumption and new capital goods depend only on the per capita stock of capital and the relative price of capital.

$$(1.4) \quad q_C = q_C(k, p_k)$$

$$(1.5) \quad q_I = q_I(k, p_k)$$

These supply relations are pictured in Figure 1 where the production possibilities curve and the price line corresponding to p_{k0} are shown.

The per capita stock of capital, k , determines the production possibility locus, and price of capital determines the allocation of resources between the production of capital and consumption goods. •

An increase in the price of capital increases output in the investment goods sector. From Figure 1, we see that as p_k increases from p_{k0} to p_{k1} , output of investment goods increases from q_{I0} to

³ Here, and for the rest of this paper, we neglect cases of specialization. In fact, the conclusions presented later in this paper carry through under such conditions. In the diagrams we use, \bar{p}_k as the price at which specialization to investment goods takes place, and \underline{p}_k as the consumption goods specialization price.

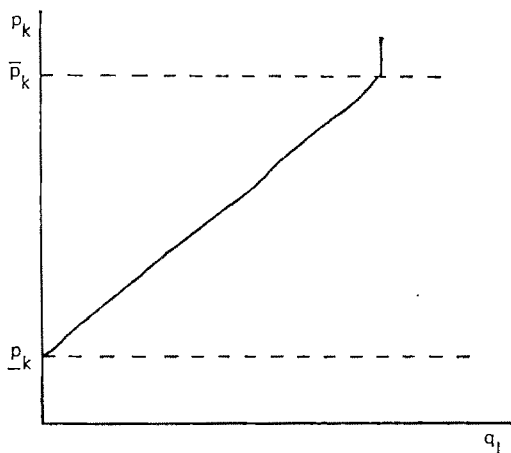


FIGURE 2

q_i . Hence

$$(1.6) \quad \frac{\partial q_c}{\partial p_k} < 0 \quad \frac{\partial q_i}{\partial p_k} > 0$$

These relationships are fundamental to the theory of investment: for any level of the stock of capital, the output of investment goods is positively related to the price of capital. This is the supply curve for investment goods, shown in Figure 2.

Since we assume that consumption goods are capital intensive ($k_c > k_i$), by the well-known Rybczynski theorem.

$$(1.7) \quad \frac{\partial q_c}{\partial k} > 0 \quad \frac{\partial q_i}{\partial k} < 0.$$

In this example the market for capital services can always instantaneously come into equilibrium for any stock of capital and any price of capital. There is no need in this case for any notion of the "desired" as opposed to "actual" flow of capital services. Rentals will always move so that firms are content with the existing flow of services. It appears to us that the same principle carries over to cases where there are many sectors and many types of capital goods.

We have so far ignored the production activities of the government. We assume

that the government hires the services of capital and labor in the market, paying the going market rental and wage rates, and produces a public consumption good. We can think of national defense or police services as being the public consumption good produced by the government. In order not to complicate the analysis, we assume that private and public consumption goods are produced by "the same" homogeneous production functions: that is, given any relative factor prices, both commodities are produced with the same amount of capital per unit of labor.

This special assumption allows us to write the production relations for the private sector

$$(1.8) \quad q_i = q_i(k, p_k)$$

$$(1.9) \quad q_c^p = q_c(k, p_k) - e$$

where q_c^p is the production of private consumption goods and e is the production of public consumption goods. That is, we assume that private and public consumption goods are perfect substitutes in production.

II. The Demand for the Stock of Capital

The demand for capital is distinct and quite different from the demand for capital services. Capital has the dimension of a stock: so many machines or buildings. It has no time dimension. The price which clears the market for capital goods is the price of capital, and it is measured in value units per machine or building.

We note again that there is no necessary connection between ownership of capital and the use of capital services. It is possible to use capital one does not own by renting it; and possible to let out owned capital to other people to use. If we take account of this, the motive for holding capital is the stream of income which it is expected to produce in the form of rentals, implicit or actual. We want to emphasize what ap-

pear to us to be two quite different kinds of decisions: how much capital services to employ, and how much capital to own.

If we insist on income as the motive for holding capital, the demand for capital is essentially an asset demand, and the price of capital will settle so that the demand for capital as an asset equals its given supply. The fact is that very often the same firm owns and employs the same capital. When this happens, we want to divide the firm's decision into two parts, a production decision and a portfolio decision.

Momentary equilibrium in the asset markets is possible only when wealth owners are content just to hold the existing stock of capital (and supplies of other assets) at going rates of return. This instantaneous equilibrium can be reached, given the value of money, by an adjustment of the price of capital and interest rates, or given the price of capital, by an adjustment of interest rates and the value of money.⁴

There may actually be many combinations of the price of capital and the value of money that instantaneously equilibrate the asset market, and to each such combination, there corresponds an equilibrium interest rate. If either price is prevented from moving, the other may still be free to find an equilibrium level. Only in the case in which both the value of money and the price of capital are prevented from changing, will there be any reason to think of disequilibrium in the asset market. Only in this case is there any reason to allow for a difference between the desired stock of capital and the actual. If prices are flexible, either the value of money or the price of capital or both will shift to make wealth owners, or their agents the firms, content to hold the existing stock of capital at each instant.

As a specific example of this fact, con-

sider an economy with three assets: money, bonds, and capital. We can write very generally the demand functions for these as:

$$(2.1) \quad m p_m = m^d p_m \equiv L(a, q, \rho_m, \rho_b, \rho_k);$$

$$1 > \frac{\partial L}{\partial a} > 0, \quad \frac{\partial L}{\partial q} > 0, \quad \frac{\partial L}{\partial \rho_m} > 0,$$

$$\frac{\partial L}{\partial \rho_b} < 0, \quad \frac{\partial L}{\partial \rho_k} < 0$$

$$(2.2) \quad b p_b = b^d p_b \equiv H(a, q, \rho_m, \rho_b, \rho_k);$$

$$\frac{\partial H}{\partial a} > 0, \quad \frac{\partial H}{\partial q} < 0, \quad \frac{\partial H}{\partial \rho_m} < 0,$$

$$\frac{\partial H}{\partial \rho_b} > 0, \quad \frac{\partial H}{\partial \rho_k} < 0$$

$$(2.3) \quad k p_k = k^d p_k \equiv J(a, q, \rho_m, \rho_b, \rho_k);$$

$$1 > \frac{\partial J}{\partial a} > 0, \quad \frac{\partial J}{\partial q} < 0, \quad \frac{\partial J}{\partial \rho_m} < 0,$$

$$\frac{\partial J}{\partial \rho_b} < 0, \quad \frac{\partial J}{\partial \rho_k} > 0$$

with

$$(2.4) \quad a = k p_k + (b + m) p_m \equiv k p_k + g p_m \\ = k^d p_k + (b^d + m^d) p_m$$

where superscript d represents demand, m is the per capita quantity of money, b the net quantity of government bonds outstanding, the H function is the *net* demand function for bonds by the private sector, p_m the price of money in terms of consumption goods (the inverse of the price level), a the per capita value of assets, q the per capita level of income measured in consumption goods ($q_c + p_k q_r$), subscripted ρ 's stand for the rate of return on the respective assets, and g is total government debt, including the money stock, outstanding.⁵

⁵ In writing the per capita demand for assets in (2.1) to (2.3), we are ignoring distribution effects, that is, we assume that aggregate portfolio decisions are indepen-

⁴ Cf. the earlier quotation from Keynes.

Equation (2.4) is the wealth constraint; at any instant the per capita value of assets demanded must be equal to the per capita value of assets held.

The value of assets held, a , enters the demand functions since it is the wealth constraint. We make the assumption that the marginal propensities to increase holdings of money and capital out of any increase in wealth are positive but do not exceed unity, while net holdings of bonds may either rise or fall as wealth increases.

Income enters the demand functions as a measure of the transactions demand for money. Cash balances yield a return in kind if payments and receipts do not exactly match for the average wealth holder, or if there is uncertainty about the timing of payments and a cost to switching from cash to bonds. Many measures of this return have been proposed, such as the aggregate value of transactions, or disposable income. We choose to measure it by real income in terms of consumption units, $q = q_C(k, p_k) + p_k q_I(k, p_k)$. Real income rises with the price of capital.⁶

We assume that an increase in the level of income increases the demand for money. But at any given level of wealth and rates of return on assets, an increase in the demand for one asset must involve a decrease in the demand for at least one other asset. In fact, we assume that neither the demand for bonds nor that for capital increases when the level of income rises.

The remaining important variables are rates of return. One dollar invested in real capital yields a return of ρ_k which has two components. The first is $r(p_k)/p_k$, the rentals rate per unit of value of capital, equal to the marginal product of capital in

the investment goods sector, which is the income wealth owners obtain from renting a dollar's worth of their capital to business firms. As we have already shown above, the rentals rate per unit of value of capital depends only on the price of capital goods; a rise in the price, by shifting resources to the labor-intensive sector, increases the capital intensity in both the consumption and the investment goods sectors and lowers $r(p_k)/p_k$. The second component of the rate of return to capital is the expected capital gain on the unit of capital, equal to the rate at which the consumption price of capital is expected to increase over time, π_k . It then follows that we can write the rate of return to capital as: $\rho_k = r(p_k)/p_k + \pi_k$.⁷

We assume that money holdings earn no interest. There still is, however, the possibility of capital gains and losses on money due to changes in the price level. Since we take consumption goods to be the numeraire, we work with the consumption goods price of money, p_m . This is equal to the amount of goods a single unit of money will buy, and is the inverse of p , the price level. We call the expected rate of change in p_m , π_m , and this is equal to the negative of the expected rate of inflation. An increase in the expected rate of inflation means a *fall* in π_m ; money will be losing value faster. The rate of return to money, ρ_m , is just equal to π_m .

To simplify the analysis, we assume that bonds have a fixed demand money price and a variable interest rate, like a savings account or a call loan. Bonds may be issued by the government or by individuals and these two instruments are assumed to be perfect substitutes. We measure the per capita quantity of bonds, b , in money units

dent of the distribution of wealth and income in the economy.

⁶ To see this, examine Figure 1; income in consumption units is the intersection of the price line with the vertical axis, and this increases when p_k rises from p_{k1} to p_{k2} .

⁷ Where there are positive transactions costs, the quantities demanded will generally depend not only on the current rental but also on the expected future path of this variable; in order to simplify the analysis, we include only the current value of this variable in the demand functions.

since a bond can always be turned into one unit of money. The interest rate, i , is determined in the market so that private individuals are content to hold the net amount of bonds the government has issued. The net quantity demanded may be negative, since at low interest rates on bonds, private individuals may want to borrow from the government to hold money or capital. Since the value of a bond is fixed in money terms, changes in p_m will also give rise to real capital gains or losses on bonds. The rate of return to bonds is $\rho_b = i + \pi_m$, the interest rate plus the expected real capital gain from holding the bond.

We also assume that the three assets are gross substitutes; that is, that an increase in the rate of return on the one of them raises the quantity demanded of this asset while it lowers the amount asset holders want to invest in the other two assets.⁸

If all returns were perfectly certain, wealth owners would hold both real capital and bonds only if they had the same rate of return. The two assets would be perfect substitutes and market equilibrium would require $\rho_k = \rho_b$, which in the absence of expected capital or losses reduces to $r(p)/p_k = i$, the Wicksellian equality of the natural and the market rate. In our model we do not assume that returns are perfectly certain, so that wealth owners, who are assumed to be risk averters, will in general diversify their portfolios.

Equilibrium in each of the three markets requires that the quantities demanded and supplied of each of the assets be equal, that is, $m = m^d$, $b = b^d$, and $k = k^d$. But, from (2.4) it can be seen that if the markets for any two of the assets are in equilibrium, then the market for the third will also be in equilibrium. We can, therefore, work with any two of the markets, and we choose to exclude the bond market.

⁸ The assumption of gross substitutability is made by James Tobin.

As of any instant, the expected rates of change in the prices of capital and money are given, as are the stock of capital, the stock of debt and, from the viewpoint of the private sector, the composition of the debt as between money and bonds. We call the ratio of total debt to money supply, x where

$$(2.5) \quad x = g/m$$

At any instant of time the government can change x by open-market operations, altering m while keeping g constant. We will discuss the determination of the important variables, π_k and π_m when we examine the problem of equilibrium over time.

Using (2.1) and (2.3) we can find the pairs of (p_m, p_k) that equilibrate the assets markets for any supplies of money, bonds, and capital. This is the "aa" schedule of Figure 3, and it is in general upward sloping.⁹ As the value of money, p_m , rises, total wealth rises thus increasing the demand for capital; to offset this excess demand, the rate of return on capital ρ_k must fall, and so p_k must rise until wealth holders are again content to hold the existing capital stock.

Note that a determinate interest rate corresponds to each point on aa, though we

⁹ Differentiation of (2.1) and (2.3) using numerical subscripts to indicate the partial derivatives with respect to the i th argument yields

$$\left. \frac{dp_k}{dp_m} \right|_{aa} = \frac{\begin{vmatrix} m - L_{1g} & L_4 \\ -J_{1g} & J_4 \end{vmatrix}}{\begin{vmatrix} L_1 k + L_2 \frac{\partial q}{\partial p_k} + L_4 \frac{\partial \rho_k}{\partial p_k} & L_4 \\ (J_1 - 1)k + J_2 \frac{\partial q}{\partial p_k} + J_3 \frac{\partial \rho_k}{\partial p_k} & J_4 \end{vmatrix}}.$$

The denominator is unambiguously negative while the numerator is definitely negative if $m > L_{1g}$. Rewriting this condition as $[(L_1 a/m)g/a - 1]m < 0$, we notice that the bracketed part of the first term is the wealth elasticity of the demand for money. Since there is no empirical evidence to suggest that the wealth elasticity is even close the inverse of g/a , which in the United States is around 6, we are justified in assuming that the inequality holds. Then "aa" is upward sloping.

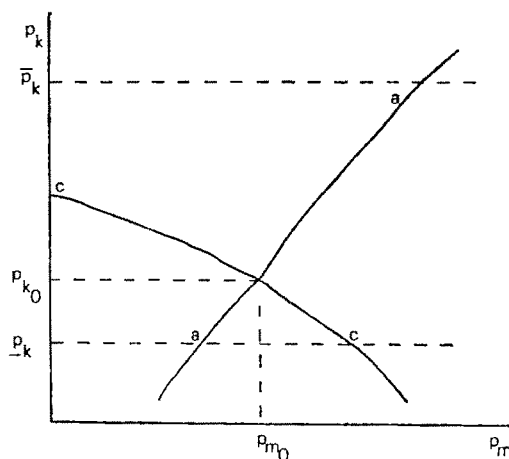


FIGURE 3

cannot in general say how the interest rate varies as we move along aa . Consider a movement up the aa schedule; the increase in p_m creates excess demand for capital and under reasonable conditions (the condition on the wealth elasticity of demand for money in footnote 9), it also creates an excess supply of money: if the increase in p_k which equilibrates the capital market still leaves excess supply in the money market, the interest rate will have to fall to equilibrate the money market (the change in the interest rate will also affect the capital market) and vice versa.

III. The Consumption Goods Market and Instantaneous Equilibrium

For the whole system to be in equilibrium at any instant, the pair (p_m, p_k) chosen for the asset market must also equilibrate the demand and supply for consumption goods. We now come to the question of the determinants of the aggregate demand for consumption goods. A satisfactory theoretical answer to this question would describe the decision which the consumer makes in dividing his disposable income among the alternatives facing him: he can buy consumption goods or save by purchasing capital, bonds, or

money. At the present time, this decision is well understood only on the assumption of perfect certainty about the future course of prices and rates of return, although recent unpublished work has extended this understanding to the case where asset holders have given subjective notions of the distribution of returns from assets. This is an assumption we do not wish to make in this paper. At this stage, we have to confess our inability to provide a satisfactory answer, and offer instead a plausible, empirically acceptable demand function without supplying any theoretical derivation of it.

We assume that consumption demand is a function of disposable income and wealth. This general formulation includes as a special case consumption functions without wealth, so that our main lines of argument do not depend on the presence of wealth affects on consumption. We ignore variables such as rates of return which may influence consumption. Substantial rate of return influences on consumption may upset some of our findings; at the present time, though, there appears to be little empirical evidence that rate of return influences are important in comparison to the factors we do include. We write our consumption function as:¹⁰

$$(3.1) \quad c^d = c(a, y) \quad \frac{\partial c}{\partial a} > 0, \quad \frac{\partial c}{\partial y} > 0$$

where y is disposable income.

Disposable income per capita is the sum of the following components: per capita factor earnings, which are equal to the per capita value of output measured in consumption goods; the value of net government taxes and transfers (including interest on bonds), which is equal to the real value of the per capita deficit, dp_m less real

¹⁰ We thus assume that private and public consumption goods are independent goods in consumption; our earlier assumption was that they were perfect substitutes in production.

government expenditure e ; and expected capital gains or losses on existing assets from changes in prices.

$$(3.2) \quad y = qc(k, p_k) + p_k q_I(k, p_k) \\ + (dp_m - e) + \pi_m g p_m + \pi_k k p_k$$

Thus for the consumption goods market to clear, we have

$$(3.3) \quad c^d = q_c - e = c(a, y) \\ = c\{a, [q_0 + p_k q_I + (d + \pi_m g)p_m \\ + \pi_k k p_k - e]\} = c^d$$

From (3.3) we obtain a "cc" schedule of pairs (p_m, p_k) which clear the consumption goods market. An increase in p_k will produce excess demand for consumption goods on three counts: first, it reduces production of these goods; second, it increases the value of wealth and the wealth effect leads to an increase in the demand for consumption goods; and third, it increases income measured in consumption goods. On the other hand, if π_k is negative, an increase in p_k would reduce demand by reducing the value of expected gains, or by increasing the amount of capital which has to be devoted to maintaining the stock of wealth. We shall assume that the first three factors outweigh this fourth effect if π_k is negative.¹¹ Of course, if π_k is positive, there is no ambiguity. An increase in p_m tends to increase consumption demand by increasing the value of wealth, and also tends to increase demand if $(d + \pi_m g)$ is positive.¹² Thus if $(d + \pi_m g)$ is positive, an increase in p_m definitely produces excess demand in the consumption goods market. In this case, an

increase in p_k would have to be accompanied by a fall in p_m to maintain equilibrium in the consumption goods market. The resultant cc schedule is shown in Figure 3.

The intersection of the aa and cc schedules gives the prices of capital and money, p_{k_0} and p_{m_0} , which equilibrate the assets and consumption goods markets. A determinate interest rate corresponds to the pair (p_{k_0}, p_{m_0}) so that the instantaneous equilibrium of the economy is fully determined. The equilibrium price of capital also determines the flow supply of investment goods. Every market is in equilibrium, there is no divergence of a desired stock of capital from the existing one, but profit maximizing suppliers of investment goods will generally be producing a flow of new investment. The higher the equilibrium price of capital, the larger this supply will be.

IV. Government Policy

The government can use its policy variables to influence the positions of the aa and cc schedules and determine the equilibrium prices of capital and money and the output of investment goods. Monetary policy operates through open market operations, changing the composition of the outstanding government debt. An open market purchase, for instance, increases the supply of money and reduces that of bonds. A pure fiscal policy would be that of the "marginally balanced budget" in which the deficit is kept constant and net taxes and expenditures are varied; we call this "pure" because it leads to no changes in the supply of debt over time compared with what that stock would otherwise have been. A deficit financed fiscal policy is one which involves changes in the deficit and so over time produces a different debt from that which would otherwise have prevailed.

It is easy to see that an open market purchase, for instance, raises the aa sched-

¹¹ In fact, in the steady state p_k will be constant so that we would expect π_k to be equal to zero. Then the possible arising from the π_k term disappears.

¹² The effects of an increase in p_m are ambiguous in the case where $(d + \pi_m g)$ is negative. We do not analyze this case here since the typical long-run behavior of an economy with a positive outstanding stock of government debt requires that $(d + \pi_m g)$ be positive to maintain a positive real debt over time

ule and leaves the cc unchanged under our assumptions, which leads to a higher p_k and a lower p_m . Expansionary fiscal policy of any kind moves the cc downward and leaves the aa unchanged, lowering both p_k and p_m in equilibrium. Since the interest rate may move either up or down along the aa schedule, there may be no consistent relationship between the level of investment and the interest rate, even *ex post*.

V. Stock-Flow Equilibria

Once the producers' equilibrium level of output in the investment goods sector is determined, it is natural to ask whether wealth owners will absorb this real capital into their portfolios at the equilibrium prices and interest rate. We can ask the same question about the government deficit, which increases the outstanding stock of government debt. The fact that actual and desired stocks of capital, money, and bonds are equal at certain prices does not guarantee that people will be content to absorb any given increases in the stocks, even though the demand and supply of consumption goods are equal. If people are not content to absorb the given additions, how will the new supplies of capital and other assets find room in private portfolios?

To begin with, we note that the private sector income budget constraint requires that private disposable income equal the private demand for consumption goods plus the value of desired additions to asset holdings, at current prices.

$$(5.1) \quad q + v p_m = c^d + \frac{DK^d}{N} p_k + \frac{DG^d}{N} p_m^{12}$$

where $G = Ng$ and v is the per capita nominal net transfers to the private sector including interest on government bonds.

The budget deficit is in turn equal to net

¹² We use the operator D to denote the time derivative of a variable. The variable d denotes the per capita deficit. N is the labor force.

transfers plus government expenditures,

$$(5.2) \quad d p_m = v p_m + e$$

and we know that

$$(5.3) \quad q = q_I p_k + q_C = \frac{DK^s}{N} p_k + q_C$$

Using these two facts in (5.1) we get

$$(5.4) \quad \begin{aligned} & \frac{DK^s}{N} p_k + q_C - e + d p_m \\ &= c^d + \frac{DG^d}{N} p_m + \frac{DK^d}{N} p_k \end{aligned}$$

Since the deficit is equal to the rate of increase in the government debt, we can write (5.4) as

$$(5.5) \quad \begin{aligned} & \frac{DK^s}{N} p_k + \frac{DG^s}{N} p_m + q_C - e \\ &= c^d + \frac{DK^d}{N} p_k + \frac{DG^d}{N} p_m \end{aligned}$$

But when the consumption market clears, $q_C - e$ is equal to c^d , so that we get the equality of total supply of new assets and total desired acquisitions.

$$(5.6) \quad \begin{aligned} & \frac{DK^s}{N} p_k + \frac{DG^s}{N} p_m = \\ & \frac{DK^d}{N} p_k + \frac{DG^d}{N} p_m \end{aligned}$$

If we add the capital gains terms to both sides of (5.1), we see that consumption market equilibrium implies that at the equilibrium prices desired and actual saving are equal. Rewriting (5.6) we have

$$(5.7) \quad \begin{aligned} & p_k \left[q_I - \frac{DK^d}{N} \right] \\ &= p_m \left[\frac{DG^d}{N} - d \right] \\ &= p_m \left[\frac{DG^d}{N} - \frac{DG^s}{N} \right] \end{aligned}$$

This indicates that it is possible for individuals to be content to hold existing stocks of capital and debt, to purchase the quantities of consumption goods desired, and therefore to be accumulating in total the value of assets they wish to accumulate, but at the current equilibrium prices to desire to add to their stocks in proportions which are different from the rates at which these stocks are being supplied. From (5.7) we see that if the government is increasing the supply of debt more rapidly than wealth holders wish to accumulate it at the current prices, wealth holders will be accumulating capital more slowly than they wish to: a flow excess supply of debt is accompanied by a flow excess demand for capital.

In fact, even if $DG^d/N = d$, so that wealth holders are accumulating just the amounts of capital and debt that they desire, they may not be accumulating debt in the desired proportions. We have

$$\frac{DG^d}{N} = \frac{DM^d}{N} + \frac{DB^d}{N}$$

and

$$d = \frac{DG^s}{N} = \frac{DM^s}{N} + \frac{DB^s}{N}$$

and the equality of DG^d/N and d does not imply that the actual rate of change of the quantities of money and bonds are equal to desired rates of change.

The results of any difference between actual and desired rates of change of asset holdings will clearly be changes in the price of capital (and its rate of return), the price of money, and the interest rate over time. If we insist on a complete dynamic model in which the asset markets are always in equilibrium, the supply of capital at every instant is the integral of past investment, and the supply of debt is the integral of past deficits, then the three variables p_k , p_m and i must follow paths which allow for the voluntary absorption of new capi-

tal and new debt. Prices may be changing, but at every instant all markets are in equilibrium and there is a determinate rate of investment.

The growth of portfolios and the absorption of capital and debt by saving are dynamic processes which can be studied only through time and which generally involve changes in the prices of capital and money and the interest rate. Accordingly, we must now proceed to a full dynamic analysis of this economy.

VI. *Equilibrium Through Time and Expectations*

We have proposed a theory of the determination of the instantaneous rate of investment which separates firm decisions into three parts: a producer's decision as to the amount of capital to employ; a portfolio decision as to the amount of capital to own; and a supply decision as to the rate of production of capital goods. This thorough-going equilibrium analysis has brought us to the paradox that the flow supplies and demands of assets may not be equal at the current equilibrium prices.

Before we discuss the solution of this paradox, one example may be helpful: the housing market. At any instant of time, the services of the housing stock are offered inelastically and the rental to housing is determined by the demand for these services; at the same time the existing stock of houses must be held in wealth owners' portfolios and the price for houses will be that price which just makes wealth owners content to hold the existing stock. Builders make their decision to supply new houses on the basis of the going market price. As these houses come onto the market, they will be absorbed into portfolios. If, however, at the existing price, the rate of change of the stock demand is not equal to the flow supply, the price of houses will have to change over time to accommodate these increases in the stock.

The moral of this example is that the

price of capital and/or the value of money will have to change through time as we move through a succession of instantaneous equilibria. Let us focus for a moment on the idea of a succession of instantaneous equilibria. These will form a path for all the variables in the system. Is any path which satisfies the instantaneous equilibria at every moment admissible? Clearly not. There are important restrictions on the way the instantaneous equilibria fit together. In a continuous time model, these are restrictions on the derivatives of the variables.

For example, the capital stock at any instant is the integral of past instantaneously determined rates of investment. Likewise, the stock of outstanding debt is the integral of past deficits. These restrictions turn the instantaneous equilibria into a system of differential equations.

At any instant the rate of change of the per capita stock of capital is equal to the per capita output of investment goods minus the amount of investment needed to provide individuals entering the economy with the existing per capita level of capital.

$$(6.1) \quad Dk = q(k, p_k) - nk$$

where n is the rate of increase of the labor force.

Similarly, the rate of change of the per capita government debt is equal to the deficit minus the amount of new debt needed to maintain a constant amount of debt per capita

$$(6.2) \quad Dg = d - ng$$

Together with (2.1), (2.3), (2.5), and (3.3), these constitute a system of six equations in eleven unknowns, k , g , m , x , p_k , i , p_m , π_k , π_m , d , and e . Obviously such a system has many possible solutions; it is underdetermined. Any paths for the eleven variables which satisfy the six equations are admissible. On each such path p_k and p_m will be changing so that equilibrium in

the assets and commodity markets is always achieved. The paradox of the last section is partially resolved. There is no need for desired and actual flows of assets to be equal at current equilibrium prices if asset prices are free to move.

But they are not free to move arbitrarily because the actual rates of change in p_m and p_k will influence the expected rates π_m and π_k . There are three popular models of this process, "static expectations," "adaptive expectations," and "perfect foresight."

Perfect foresight requires $\pi_m = Dp_m/p_m$ and $\pi_k = Dp_k/p_k$ at each moment on the whole path. If in addition we were to specify equations describing the time paths of the policy variables d , x , and e , we would have a complete system of eleven equations in eleven unknowns, determining fully the time path of the economy.

While the assumption of perfect foresight in problems of intertemporal economics seems to be a natural extension of the assumption of perfect information usually made in static equilibrium models, it imposes severe restrictions on the system and its path. First, it rules out all notion of uncertainty and portfolio diversification; second, it requires assumptions about information which are unlikely to be met in reality; and third, and most important, it leaves almost no room for discretionary government policy. A strong condition on government policy which is consistent with perfect foresight is that future government policy is known and therefore unalterable. In this case, the complete paths for policy variables are determined from time zero, and no further change can occur without violating the perfect foresight assumption.

Government discretionary policies are consistent with perfect foresight when they do not induce regrets in economic agents. But even with this weaker condition, most of the policies considered in the analysis of Section V, which seem to be the types of

policy most government use to control aggregate demand and its composition, would be excluded because they do induce regrets.

In the other models of expectation formation, where actual price changes influence expected price changes either not at all or with a lag, there will not be an intertemporal competitive equilibrium, only a succession of instantaneous equilibria based on possible wrong guesses about the future. There is a larger sphere for discretionary government policy and the set of solution paths becomes wider. There will also be regrets which imply that individuals and firms are not in intertemporal equilibrium. They will be in instantaneous equilibrium given their imperfect information about the future.

It seems to us, then, that the requirement of intertemporal competitive equilibrium is very strong, and requires assumptions about information which are unlikely to be met in reality. There seems to be a place for a theory that allows for a lack of intertemporal equilibrium while insisting on instantaneous equilibrium.

Static expectations models, however, are too naive, particularly where prices may actually be changing in the economy. Accordingly, in the following section we analyze two models using the adaptive expectations hypothesis which allows for errors and attempts to correct these errors on the basis of newly available information.¹⁴

VII. Dynamics

In order to analyze the relationship between the prices of capital and money, the rate of interest and stocks of different assets over time, we consider two simple dynamic models. In both of them we assume that the government actively

manipulates either the composition of the debt (monetary policy) or the levels of expenditure and taxes (marginally balanced budget fiscal policy) to achieve a stable consumer price level, so that

$$(7.1) \quad p_m = p_m^*$$

To simplify matters, we assume that the government maintains constant the outstanding stock of nominal debt by fixing its deficit at the appropriate level.

$$(7.2) \quad d = ng^*$$

Given a constant price level over time, it is reasonable to assume that the expected rate of change of p_m is equal to zero:

$$(7.3) \quad \pi_m = 0$$

If monetary policy is used to stabilize the price level and fiscal policy is passive we have an additional equation

$$(7.4a) \quad e = e^*$$

while if fiscal policy is used to control the price level and monetary policy is passive, the additional equation is

$$(7.4b) \quad x = x^*$$

We have two possible models, depending on whether we use (7.4a) or (7.4b); both of them are summarized in systems of ten equations in the eleven unknowns: the six mentioned in Section VI plus (7.1) through (7.4).

In both models we are missing an equation describing the process by which wealth owners and consumers form their expectations about the rate of change in the price of capital. Since they are not assumed to have perfect foresight, they will tend to make mistakes which they will probably try to correct as new information becomes available to them. A simple model of this type is the adaptive expectations model in which the rate at which people adjust their beliefs about the rate of change of the price of capital depends on the error

¹⁴ A version of the adaptive expectations hypothesis has been used by one of the authors in an earlier growth model. See Sidrauski.

made in predicting the current rate of change:

$$(7.5) \quad D\pi_k = \beta \left[\frac{Dp_k}{p_k} - \pi_k \right]$$

A standard policy argument is that an easy monetary policy combined with a tight fiscal policy will promote growth, while a tight monetary policy together with easy fiscal policy encourages consumption at the expense of investment. The degree of ease or tightness is probably thought of in terms of the level of "the" interest rate or the general level of interest rates. We have shown, however, that there is no necessary relationship between the interest rate and the price of capital which determines the output of investment goods at each instant. We choose to define an easy monetary policy in terms of the composition of the debt: an increase in the proportion of money in the stock of outstanding debt—a fall in x —represents an easing of monetary policy. A tightening of fiscal policy is represented by a decrease in the level of government expenditures and taxes while the deficit is kept constant or by a fall in the deficit.

In this section we shall consider the policy argument outlined above: first, we examine the effects of a decrease in the debt-money ratio where a marginally balanced budget fiscal policy is used to stabilize the price level; and second, we consider the effects of an increase in government expenditures when monetary policy is used to stabilize the price level.

Stabilization of the Price Level Through a Marginally Balanced Budget

In this case, given the stock of capital inherited from the past and given the expected rate of capital gains based on the past behavior of the price of capital, equations (7.1) and (7.2) determine the price of capital and the interest rate which equilibrate the assets markets.

$$(7.6) \quad p_k = \Psi(k, \pi_k; g^* p_m^*, x_0)$$

with

$$\frac{\partial p_k}{\partial k} < 0, \quad \frac{\partial p_k}{\partial \pi_k} > 0, \quad \frac{\partial p_k}{\partial x} < 0^{15}$$

Given p_k determined in the assets markets, equation (3.3) indicates the level of government expenditure, e , consistent with equilibrium in the commodity market for p_m^* and the asset market equilibrium price of capital.

Differentiating equation (7.6) with respect to time and substituting into (7.5) we have

$$(7.7) \quad D\pi_k = \frac{\beta \left[\frac{1}{p_k} \frac{\partial \Psi}{\partial k} Dk - \pi_k \right]}{\left[1 - \beta \frac{\partial \Psi}{\partial \pi_k} \frac{1}{p_k} \right]}$$

Substituting (6.1) into (7.7) we can rewrite the basic differential equations of the model as equations (7.8) and (7.9).

¹⁵ The signs of these derivatives may be confirmed by differentiation of (2.1) and (2.3).

$$(7.8) \quad Dk = q_I[k, \Psi(k, \pi_k; g^* p_m^*, x_0)] - nk$$

$$(7.9) \quad D\pi_k = \frac{\beta \left\{ \left(\frac{\partial \Psi}{\partial k} \frac{1}{p_k} \right) (q_I[k, \Psi(k, \pi_k; g^* p_m^*, x_0)] - nk) - \pi_k \right\}}{\left[1 - \beta \frac{\partial \Psi}{\partial \pi_k} \frac{1}{p_k} \right]}$$

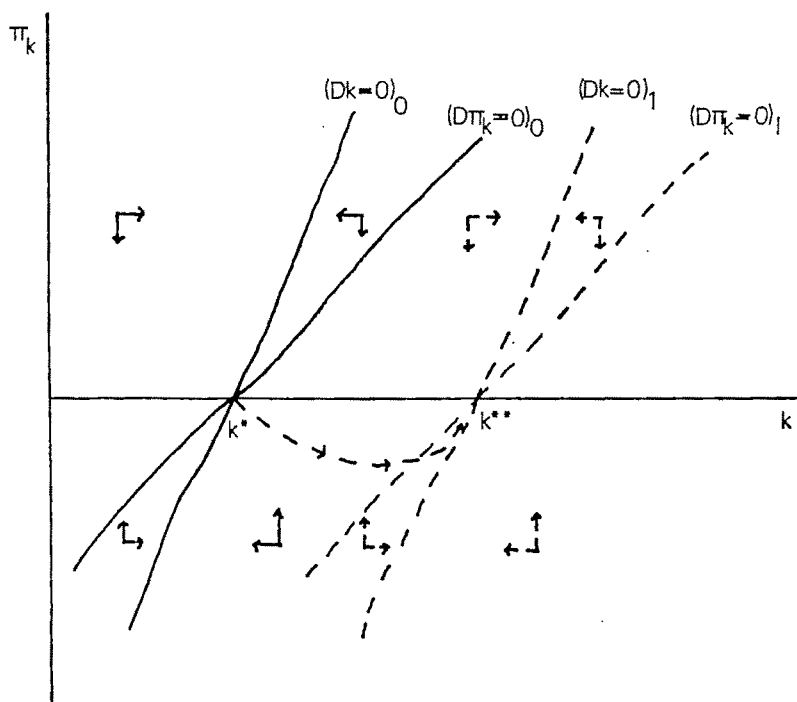


FIGURE 4

The $Dk=0$ and $D\pi_k=0$ lines in Figure 4 indicate the pairs (k, π_k) that make Dk and $D\pi_k$ respectively equal to zero. Figure 4 corresponds to the case in which the denominator in equation (7.9) is positive; that is, the case in which the lag in the adjustment of expectations is sufficiently large to avoid the perpetuation of runaway boom in the assets markets.¹⁶

It is important to note that in this model, the more slowly that wealth owners adjust their mistaken expectations, the more likely it is that the long-run balanced growth path is stable.^{17,18}

We may now use the model to consider

¹⁶ A similar stability condition is also to be found in Philip Cagan and Sidrauski.

¹⁷ It can be proved that in this case under rather weak assumptions the equilibrium exists and if it exists it is unique.

¹⁸ Problems of stability in the assets markets arising from a rapid adjustment of expectations lie behind the recent controversy on models with heterogeneous capital goods. See F. H. Hahn, and Karl Shell and Joseph E. Stiglitz.

the effects of an easing of monetary policy: that is, an increase in the proportion of money in the debt, equivalent to a fall in x . We examine only the stable case shown in Figure 4. An open market purchase which produces a fall in x leads to an increase in the price of capital which clears the assets markets. In terms of the diagrams, this shifts both the $(Dk=0)$ and $(D\pi_k=0)$ schedules to the right since the output of investment goods will now be higher for any pair (k, π_k) and a higher capital stock will be needed to absorb the additional output of investment goods in the steady state. The capital stock increases continually to its new higher level so that the overall rate of growth will be higher in the period of disequilibrium than in the steady state. The expected rate of increase in the price of capital initially falls below zero as the accumulation of capital forces p_k down after its first upward jump. The fall in π_k depresses p_k even further for awhile, but

$$(7.11) \quad Dk = q_I[k, \Phi(k, \pi_k; g^* p_m^*, d p_m^*, e)] - nk$$

$$(7.12) \quad D\pi_k = \frac{\beta \left\{ \left(\frac{\partial \Phi}{\partial k} \cdot \frac{1}{p_k} \right) (q_I[k, \Phi(k, \pi_k; g^* p_m^*, d p_m^*, e)] - nk) - \pi_k \right\}}{\left[1 - \beta \frac{\partial \Phi}{\partial \pi_k} \frac{1}{p_k} \right]}$$

the depressing effect of p_k of the rise in k diminishes as k approaches its steady state value and Dk approaches zero. In the end π_k , following the actual rate of change in p_k , moves back to zero.

What is required of fiscal policy in order to stabilize the price level following the rise in the price of capital and the subsequent accumulation of capital? For the reasons outlined in Section III, an increase in the price of capital has an inflationary effect in the consumption goods market. To offset the effects of the initial rise in the price of capital, then, fiscal policy has to be tightened—the level of government expenditures and taxes has to be reduced. This is what the conventional accounts lead us to expect. Then as capital accumulates over time, the price of capital in the assets markets begins to fall from its new level so that fiscal policy can be eased on this account. However, the accumulation of capital also affects the equilibrium of the consumption goods market though the effects are ambiguous since an increase in the capital stock increases supply and increases demand through wealth and income effects. After the initial tightening the direction of fiscal policy is uncertain.

Stabilization of the Price Level Through Monetary Policy

In this case, given the stocks of government debt and capital inherited from the past, π_k and p_m^* and the government's policy parameters e and d , the price of capital is determined in the consumption goods market (3.3).

$$(7.10) \quad p_k = \Phi(k, \pi_k; g^* p_m^*, d p_m^*, e),$$

$$\frac{\partial \Phi}{\partial k} > 0, \quad \frac{\partial \Phi}{\partial \pi_k} < 0, \quad \frac{\partial \Phi}{\partial e} < 0$$

Given p_k determined in the consumption goods market, the government has to vary the composition of the debt in such a way as to ensure that the p_k determined by (7.10) together with p_m^* clear the assets markets.

Differentiating (7.10) now with respect to time and substituting into (7.5) and using (6.1), we obtain equations (7.11) and (7.12).

There are now three possible stable configurations of the ($Dk=0$) and ($D\pi_k=0$) loci depending first on whether an increase in the capital stock increases or decreases the equilibrium price of capital in the consumption goods market; and second on whether an increase in the expected rate of change of the price of capital increases or decreases the rate at which that expected rate is changing (i.e., whether or not $\partial D\pi_k / \partial \pi_k \geq 0$). It is a necessary condition for stability that an increase in the capital stock decrease the rate of change of the capital stock (i.e. $\partial Dk / \partial k < 0$). This condition need not always be met in practice since an increase in the capital stock may *increase* the equilibrium price of capital in the consumption goods market and in this way increase the output of investment goods. There are two factors working in the opposite direction. First, the increase in k by itself lowers q_I . Second, the increase in x raises k and lowers Dk . If an increase

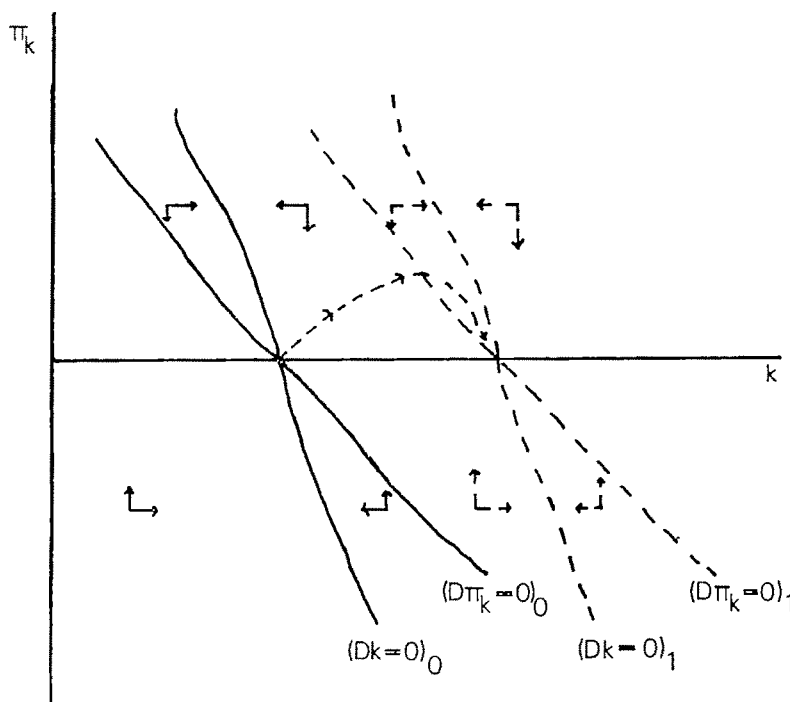


FIGURE 5

in k produces so great a rise in p_k through the consumption market that it overwhelms these negative factors, the system will be unstable.

We show the three possible stable equilibria in Figure 5, 6, and 7. In each diagram the $(Dk=0)$ locus is downward sloping since an increase in the capital stock reduces the rate of change of the capital stock, so that a decrease in π_k —which increases the price of capital at which the consumption market clears—is required to offset this effect. The horizontal arrows indicate that $\partial Dk/\partial k < 0$.

While the $Dk=0$ schedule must slope downward near a stable equilibrium, there are two possibilities for the $D\pi_k=0$ schedule. First, there is the case where $\partial D\pi_k/\partial \pi_k < 0$. In this case the $D\pi_k=0$ schedule must have a higher slope than the $Dk=0$ schedule near a stable equilibrium, as illustrated in Figures 5 and 6. There are

no oscillations possible in the former case. Second, is the case where $\partial D\pi_k/\partial \pi_k > 0$. The stable equilibria on this assumption are like the one shown in Figure 7, and the system can produce cycles.

We are now ready to examine the effects of a tightening of fiscal policy, that is, a decrease in the level of government expenditures and taxes. A decrease in the level of government expenditures with a constant deficit is deflationary in the consumption goods market so that the price of capital which clears that market rises as for any pair (k, π_k) . Around a position of stable equilibrium this will shift the $(Dk=0)$ and $(D\pi_k=0)$ schedules to the right, increasing the equilibrium stock of capital since a higher capital stock is now needed to absorb the higher output of investment goods in the steady state. If the economy were initially at a position of stable long-run equilibrium, the capital

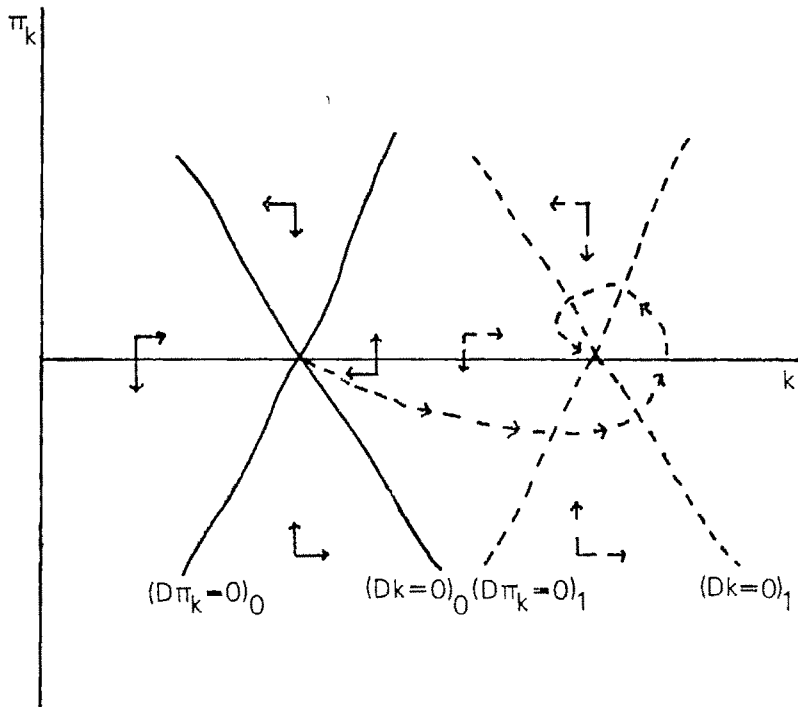


FIGURE 6

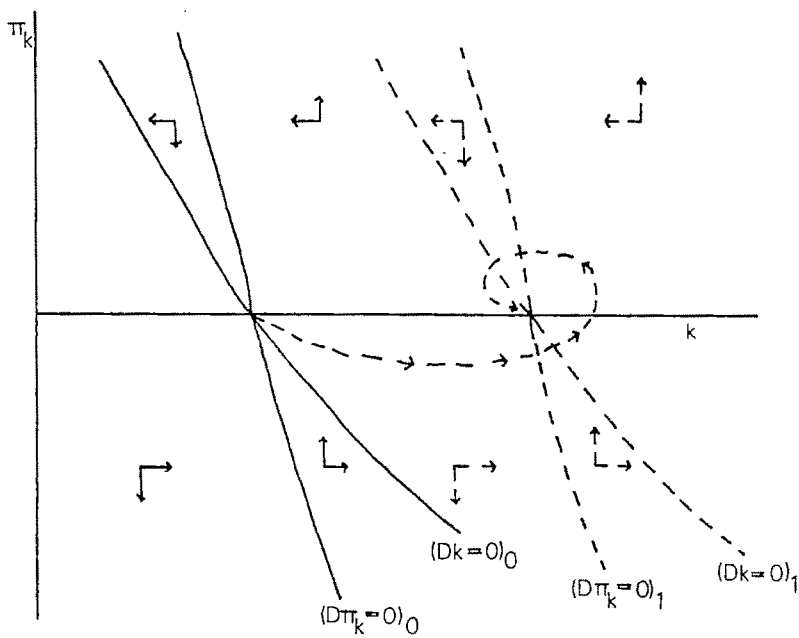


FIGURE 7

stock increases to its new equilibrium level. As can be seen from the diagrams, the movement to the new equilibrium capital stock may involve cycles.

The long-run effect of the tighter fiscal policy is a higher stock of capital, just as the long-run effect of easier money in our previous examples was a higher stock of capital. At first the accompanying monetary policy must be easy to prevent changes in the price level, but the continuing changes necessary in monetary policy depend on the effect of the growing capital stock on the consumption market. If as the capital stock increases it raises the consumption market equilibrium p_k , monetary policy must always be getting easier, to achieve the necessary p_k with the larger capital stock. If a rise in the capital stock lowers the consumption market equilibrium p_k , x may have to move in different directions at different times.

These two experiments partly confirm the policy argument with which we began this section. In the short run, tightening fiscal policy and easing monetary policy will raise the rate of growth. In the long run, however, one of these policies may have to be reversed as capital accumulates to maintain a stable price level.

VIII. Conclusion

In this paper we have attempted to tie together an equilibrium theory of investment, Keynesian stabilization policies, and the neoclassical two-sector model of economic growth to get a coherent view of the behavior of modern indirectly controlled economies. The particular models of government behavior and expectations formation we used in the last section do not exhaust the potential of models based on these ideas. We encourage the reader to "roll his own" model out of the fixings we have prepared.

REFERENCES

- P. Cagan, "The Monetary Dynamics of Hyperinflation," in M. Friedman, ed. *Studies in the Quantity Theory of Money*, Chicago 1956, 25-120.
- R. Eisner and R. H. Strotz, "Determinants of Business Investment," in D. B. Suits et al., *Impacts of Monetary Policy*, Englewood Cliffs 1963.
- J. P. Gould, "Market Value and the Theory of Investment of the Firm," *Amer. Econ. Rev.*, Sept. 1967, 57, 910-13.
- T. Haavelmo, *A Study in the Theory of Investment*, Chicago 1960.
- F. H. Hahn, "Equilibrium Dynamics with Heterogeneous Capital Goods," *Quart. J. Econ.*, Nov. 1966, 80, 633-46.
- D. W. Jorgenson, "Capital Theory and Investment Policy," *Amer. Econ. Rev. Proc.*, May 1963, 53, 247-59.
- J. M. Keynes, "The General Theory of Employment," *Quart. J. Econ.*, Feb. 1937, 51, 209-23; reprinted in S. E. Harris, ed., *The New Economics*, New York 1948, 181-93.
- R. R. Nelson, "Full Employment Policy and Economic Growth," *Amer. Econ. Rev.*, Dec. 1966, 56, 1178-92.
- T. M. Rybczynski, "Factor Endowment and Relative Commodity Prices," *Econometrica*, Nov. 1955, 22, 336-41.
- K. Shell and J. E. Stiglitz, "The Allocation of Investment in Dynamic Economy," *Quart. J. Econ.*, Nov. 1966, 81, 592-609.
- M. Sidrauski, "Inflation and Economic Growth," *J. Polit. Econ.*, Dec. 1967, 75, 796-810.
- J. Tobin, "An Essay on the Principles of Debt Management" in W. Fellner, et al., *Fiscal and Debt Management Policies*, Englewood Cliffs 1963, pp. 143-218.
- A. Treadway, "Rational Entrepreneurial Behavior and the Dynamics of Investment," unpublished doctoral dissertation, Univ. Chicago 1967.
- J. G. Witte, "The Microfoundations of the Social Investment Function," *J. Polit. Econ.*, Oct. 1963, 71, 441-56.

Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital

By JOSEPH P. NEWHOUSE*

The private nonprofit firm has been ignored by economic theory until very recently. It was easy for economists to overlook such firms in the past because of their relative unimportance, at least for the past century and a half.¹ But presently this sector has grown to a position of importance. In 1966, nearly \$15 billion was spent on hospital care, not including expenditures on construction, research, or insurance administration charges.² Hospitals employed nearly 1.3 million people in 1963, over twice as many as "Blast Furnace and Basic Steel Products" and nearly twice as many as "Motor Vehicles and Equipment."³ Since decisions made by nonprofit institutions affect the allocation of resources, it is important that their decision-making process be understood.

In this paper a very simple model of a hospital is developed, and its implications are considered at some length. An attempt is made to justify the realism of this model, though like any model it cannot be entirely realistic. To develop the model, we will make the particularly unrealistic assumption that hospital expenses are financed by the consumer and not by a third party. We do this in the hope that this simple model may prove applicable to other nonprofit institutions where third party payments are not as important;

such as colleges and universities, the performing arts, and museums. Later, however, we remove the assumption that the consumer pays his own bill.

We are concerned with the relationship between a hospital's nonprofit status and economic efficiency. To understand that relationship, we must postulate a maximand for the hospital decision maker. The first element in the maximand is quantity of services provided. Hospital services seem to be desirable in some ethical sense which justifies the claim that consumers have a "right" to medical care. It is the basis for granting hospitals certain tax and other legal privileges, such as exemption from the Robinson-Patman and Taft-Hartley Acts, and it seems to be the *raison d'être* for philanthropy. Apparently it is felt that the public is better off if it consumes more hospital services (as well as more of the services provided by other nonprofit institutions).⁴ No doubt the public would be better off (and think itself better off) if it consumed more of many different kinds of products. Yet that is irrelevant to a decision maker at a hospital. If his institution exists for a social purpose and because of that can ask for gifts and tax privileges, he is likely to be concerned about the quantity of the service provided. Therefore, we take quantity as one element in the decision maker's maximand. Maximizing quantity implies (on the assumption of a downward-sloping demand curve) keeping price as low as pos-

* The author is a member of the economics department of The RAND Corporation.

¹ There is some evidence that the nonprofit corporation was the dominant form of business organization in the colonial period. See Joseph S. Davis.

² See H. M. Somers and A. R. Somers (p. 43).

³ See U.S. Department of Labor, B.L.S.

⁴ Apparently the public is better off if it consumes less of some other services.

sible. It may also involve price discrimination such as charity care (or scholarships based on need).

To understand why the second element, quality, belongs in the maximand, it is necessary to examine the locus of decision making in a hospital. One characteristic of nonprofit hospitals is that usually control formally resides in a board of trustees or similar group. The board in turn appoints an administrator who is in charge of day-to-day decisions. The medical staff also generally exerts influence over resource allocation decisions. It is important to know what incentives these various parties face in making decisions regarding resource allocation.

If the administrator is not to make a "profit," his performance cannot be judged by the profit criterion. Therefore, his salary and promotional chances must be a function of some other variable or variables. It seems plausible to assume that the prestige of the institution is prominent among these other variables. The trustees, insofar as they participate in the decision-making process, may also be influenced by this variable. Prestige, in turn, is affected by the size of the institution, but probably even more by the quality of the product produced.

There may be other reasons why the trustees and the administrator would give weight to both quantity and quality.⁵ There may be a pursuit of status quite independent of any managerial reward. There may be a desire to serve society independent of the desire to preserve existing tax and legal privileges. There may be a desire to show professional excellence or technical virtuosity by stressing quality. In short, while we have derived a maximand based on quantity and quality by considering the self-interest of the administrator (and trustees) narrowly defined, such a maximand is consistent with other motivations.

⁵ The remainder of this paragraph is based on references found in Fritz Machlup's article.

The maximand is reinforced by whatever role the medical staff may play in the decision-making process. They have a strong interest in the quality of the facilities available since it is one determinant of the quality of care they can give and of their professional standing. Further, the existing staff will find it easier to attract additional staff (and so ease their own work load) by maintaining high quality facilities. The medical staff is also interested in quantity, since each physician wants a bed available should a patient of his require hospitalization.⁶

The administrator, the trustees, and the medical staff may, of course, weigh quality and quantity considerations differently, but that need not impair the theory. We assume that some final resolution is obtained among the tastes of the administrator, the trustees, and the medical staff, so that we can speak of the tastes of the hospital decision maker.

Support for the notion that decision-makers perceive both quality and quantity to be in their maximand can be found in the hospital trade literature. One prominent hospital administrator wrote, "No one can seriously believe that the public would knowingly permit any step to be taken that would lead to the slightest sort of deterioration of quality."⁷ And as Jerome Rothenberg said, "... medical ethics and collective concern for quality significantly modify free enterprise . . . In medicine we are not free to envisage saving resources by lowering the quality of care. Less than the highest quality care often represents a total waste."⁸ Another leading hospital administrator has written:

The hospital is the community trustee responsible for a large amount of the definitive medical care provided in a community. Its first responsibility is to program its service to meet community

⁶ See Carl Stevens.

⁷ See Ray E. Brown (p. 35).

⁸ (p. 676.) Also, it is not clear why a concern for quality should modify free enterprise.

needs, and its second is to conduct such services with suitable efficiency and economy, always with great sensitivity toward its responsibility and authority to meet high standards of both quality and quantity of services . . . The patient expects that hospital services will be of high quality.⁹

A voluminous study of the medical market done for the State of Michigan said: "From the community's viewpoint, the hospital is most effective . . . when it admits the greatest number of patients who need admission . . ."¹⁰

Yet quality and quantity cannot be maximized without limit. The nonprofit institution faces a budget constraint; its deficit cannot be larger than a certain amount. We therefore postulate a model of constrained quantity-quality maximization. Millard Long has informally put forth a similar hypothesis, but he does not draw any conclusions from his model.¹¹ He saw the "guiding principle" of the hospital as a

. . . desire to maximize the number of patients seen subject to several constraints. There is a financial constraint; operating deficits cannot go beyond a point specified by the sponsoring agency. Another constraint is that the quality of care should be the best possible with available equipment and personnel; hospitals seldom cut corners when doing so would reduce the quality of care. [p. 212]

Our model differs in an important respect from Long's by making quality a variable of choice rather than a constraint. The

⁹ See James P. Dixon (p. 284). The first part of this quotation ignores the vital question of how one defines community needs and whether needs are different from community demands. Nevertheless, we are interested in the light it sheds on the maximand. The administrator of a different nonprofit institution has said, "As dean of a Graduate School of Public Health, reasonably alert to the attractiveness of having the largest possible number of tuition-paying students that his facilities can accommodate . . ." James A. Crabtree, p. 1179.

¹⁰ Walter J. McNerney et al., *Hospital and Medical Economics*, quoted in Mark V. Pauly.

¹¹ W. J. Baumol and W. G. Bowen also emphasize both quality and quantity in the nonprofit area generally, but do not formalize their ideas.

existence of accreditation bodies, however, may make some minimum standard of quality necessary. Insofar as it does, a constraint does exist.¹²

How are quality and quantity measured? One can think of certain criteria which are indicative of quality: personnel/patient ratios or professional personnel/patient ratios or the availability of certain laboratory or other facilities. There are, however, several difficulties associated with the use of such criteria. First, do high personnel/patient ratios indicate high quality or merely substitution of labor for capital (or low-skilled labor for high-skilled labor) in the hospital production function?¹³ Similarly, is the availability of certain facilities also merely a substitution in the production function? Second, how may these criteria be combined in any meaningful fashion? The weight each criterion would receive in a weighted average is ambiguous, as is the meaning of any such average. Finally, there are intangibles associated with the notion of quality.

To avert these difficulties, at least in part, we assume quality to be represented by a vector of characteristics, some of which may not be quantifiable except in the sense of being present or absent. Further, we assume that the demand for the services of each institution depends upon

¹² Martin Feldstein (Ch. 7) has postulated a maximand for a hospital decision maker in the British National Health Service which is somewhat similar to ours, but used for a different purpose. Feldstein discovered that the number of cases treated in a region was more responsive to bed availability than was length of stay. He found that such behavior was consistent with a three-variable maximand, which was maximized subject to a budget constraint. The maximand included length of stay, number of cases treated, and a quality variable. Our interest, however, is in a different question, namely, the tradeoff between quality and quantity.

¹³ Feldstein (Ch. 4) finds that a modified Cobb-Douglas production function with nurses, physician beds, and other supplies as explanatory variables fits data generated by the British National Health Service rather well. In the modified form, nursing services are not substitutes for other inputs, but the other inputs are substitutes for each other.

quality as well as price. The justification for this assumption is that physicians probably prefer higher quality hospitals and so are more inclined to seek staff privileges there. Also, if the physician has multiple privileges, he may prefer working in the higher quality hospital (and given the extent of hospital insurance, he probably will not find it too difficult to convince the patient). When two quality vectors have the same cost, we assume that the hospital decision maker chooses that quality vector which maximizes quantity bought at a given price. The implication of these assumptions is that an increase in quantity demanded at each price which is brought about by an increase in quality can only be accomplished at an increased cost. Restricting ourselves to this subset of quality vectors which has the property of maximizing quantity demanded at a given price, we can associate each quality vector with a level of cost. We can then attach an arbitrary set of numbers to each quality vector, which serve as an ordinal measure of quality. The only restriction on the numbers is that they must increase as cost increases. For convenience we shall use the costs themselves as measures of quality.¹⁴

This approach averts the problem of measuring quality directly, but unfortunately measurement difficulties are not restricted to that variable. Quantity supplied and demanded, while seemingly

straightforward, also presents an analytical problem. We take it to mean the number of patient-days in the case of a hospital. (One could measure the number of students attending a medical school.) However, the "product" of a hospital may vary so much depending on the diseases its patients have that a simple measure of patient-days cannot accurately reflect the output of the hospital.¹⁵ This is an aggregation problem which is inherent in a multiple-product firm, since the treatment of each diagnosis can be seen as a distinct product.

If there are distinct demand and cost curves for various diagnoses, each must be analyzed as a separate product with any interdependencies acknowledged. To take account of subproducts and interdependencies here would complicate the analysis, but would not alter the conclusions. Therefore, we shall continue to speak of the hospital as though it were a single product firm whose physical output were unambiguously measurable.

Suppose quality is given, say, at the minimum permissible level for accreditation. This determines an average cost curve—call it AC_0 . At this quality, income and all other relevant variables except price held constant, there is a certain demand at each price which determines a demand curve, call it d_0 . Assume for now that the decision maker cannot run a deficit—all his costs must be met from revenues. Then, given that he wants to maximize the quantity of output (in physical terms) provided at that quality, and, assuming a downward sloping demand curve, he produces at the quantity q_0 for which $AR=AC$. If there is more than one point where $AR=AC$, he chooses the one associated with the largest quantity. Such an outcome is graphed in Figure 1.

Suppose that a higher quality product is available at a cost AC_1 , a cobalt radiation machine can be bought or more nurses can

¹⁴ This is another formulation of a problem which has caused some controversy in the literature, namely, how to measure quality. J. L. Nicholson has criticized Milton Gilbert for using cost rather than price to measure the contribution to welfare of a change in quality. Since cost equals price in our model, this criticism presents no problem to it; one could merely say that the assumptions imply that an increase in quality, quantity held constant, implies an equal increase in both cost and price. Our analysis is really in the same spirit as Nicholson's by proposing a criterion which relates to the consumer's preferences as revealed in the marketplace; that is, that the decision-maker is in equilibrium at the quality level which maximizes quantity bought at a given price when two quality levels have the same cost. Gilbert's measure, on the other hand, seems to be more a technological criterion.

¹⁵ See Martin Feldstein (Ch. 2).

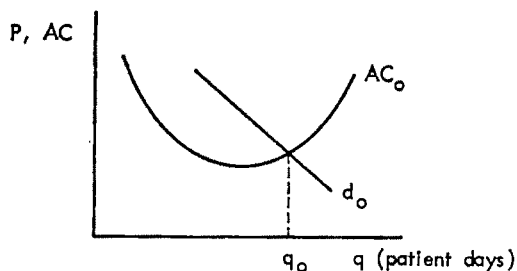


FIGURE 1

be hired. This higher quality product raises the demand curve to d_1 , so that there is a new equilibrium output q_1 as shown in Figure 2.

Depending upon the relative movement of the demand curve and cost curve and their shapes, the new equilibrium point may lie to the right or left of the old one. As the quality variable runs over its potential range of values, the locus of equilibrium quantity-quality combinations is traced out. If, after a certain point, further increases in quality shift demand up less and less while raising costs more and more, the tradeoff curve will eventually bend back, as is illustrated in Figure 3.¹⁴

The decision maker will choose the point on this tradeoff curve which yields him the highest utility. That will be where the curve is tangent to the highest attainable indifference curve, I_0 . (We assume the decision maker's indifference curves have the usual shape.) This outcome is shown at point A in Figure 3.

Suppose that the decision maker is told he may run a deficit of a certain size. Then, instead of producing the quantity in Figures 1 and 2 where $AR = AC$, the decision maker produces a sufficiently larger quantity to exhaust the subsidy given him. This shifts the tradeoff curve to the right at each quality level (or, alternatively, up at

each quantity level), thereby enabling the decision maker to attain a higher indifference curve. Suppose the deficit is not given to the decision maker, but is affected by his actions. Such would be the case if the decision maker conducted fund-raising drives whose success depended upon the effort he put forth. The fund-raising drive will be carried on until the marginal benefit to the decision-maker of shifting the tradeoff curve out by the amount the marginal dollar would do so equals the marginal cost of raising that dollar.

We are interested in using this model to assess the effect of the hospital's nonprofit status upon efficiency. First, note that this model implies least-cost production insofar as the decision maker pursues his maximization goals. Suppose that the marginal revenue product of a factor exceeded its marginal factor cost. A profit-maximizing firm would expand production to

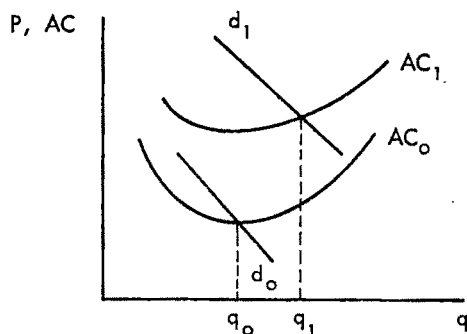


FIGURE 2

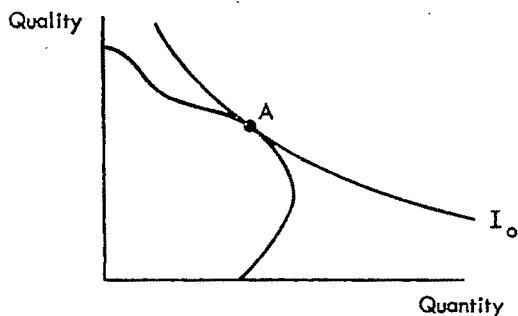


FIGURE 3

¹⁴ For a discussion of such a tradeoff curve in the context of the traditional theory of the firm and the problems posed by such a curve for the traditional theory, see Henri Theil and Jack Hirschleifer.

the point of equality, thereby achieving a socially optimal allocation of factors.¹⁷ A constrained quantity-quality maximizer will also expand production to the point of equality. He can use the "profit" the profit-maximizing firm would have gained to shift his quantity-quality tradeoff curve out and thereby reach a higher indifference curve. When the equalities hold for all factors, the tradeoff curve cannot be shifted out any more.

Even though the model implies least cost production, there are two reasons why it does not lead to an optimal outcome. (These reasons also apply to a model based on simple cost reimbursement by a third party.) The two reasons are a bias against producing lower quality products and barriers to entry resulting from nonprofit status. Up until now we have treated the hospital as if it had produced only one quality. In reality it produces several. From a normative standpoint one would desire that the hospital produce all qualities (all products) which were profitable when price equalled marginal cost, just as a profit-maximizing firm would. How likely is this outcome in the nonprofit sector? To bring out the essence of the problem, we assume all quality vectors have demand and cost functions which are independent of those associated with other quality vectors. This assumption is made for ease of exposition.

Different quality levels generate demand which can be measured in the same units of quantity, for example, patient-days. This makes it possible to derive total or aggregate quantity; that is, quantity summed over all quality levels. Will the decision maker try to maximize this figure subject to the constraint that all qualities produced at least break even? If he does, the outcome is similar to that predicted by traditional theory in a profit-maximizing

sector.¹⁸ The answer is no, he will not necessarily do so. This is a major difference between the implications of this model and those of the profit-maximizing model, and so from a positive point of view constitutes one test of the model. The reason why the decision maker does not produce all profitable qualities lies in the quality variable which is in his maximand. An example should make this clear. A hospital can produce hospital care in wards, semiprivate rooms, or private rooms, where quality increases from ward to private room. Why should the hospital produce care in a private room when the same patient-day would be counted a patient-day in a ward? Because if the patient is willing to pay the additional cost of private room care, the hospital can raise its quality level without changing its quantity level. But suppose the hospital produces only top quality care. In that case it could conceivably increase quantity by producing lower qualities, but this would lower average quality. Whether it will produce lower qualities depends on how much inferior qualities will increase quantity and on how much weight the decision-maker gives quantity relative to quality. Such a bias against producing low-quality products, even if they are demanded by a certain segment of the population, does not exist for a profit-maximizing firm.

Is such a bias observed? First, proprietary hospitals (or hospitals for profit) are thought to produce lower quality care than voluntary short-term general hospitals. There is a lower percentage of accreditation among the proprietaries than among the voluntaries. Of all the short-term hospitals listed by the American Hospital Association in 1965, 62 percent (with 87 percent of the beds) were accredited, but

¹⁷ Assuming perfect product and factor markets and no externalities.

¹⁸ Obviously this will not be the outcome in the profit-maximizing sector if the demand and cost functions for various products are not independent. But that does not affect the conclusion.

TABLE 1—TOP NURSING SKILL LEVEL AMONG FULL TIME STAFF IN SKILLED NURSING HOMES*

	Proprietary Homes (percent)	Nonprofit Homes (percent)
Registered Nurse	44	67
Licensed Practical Nurse	42	22
No RN or LPN	14	11

* Based on the 1961 National Inventory, conducted by the Public Health Service.

only 34 percent of the proprietaries (with 60 percent of the proprietary beds) were accredited.¹⁹

Second, one thinks of the description of medical care as a "Cadillac only" industry.²⁰ "The potential for 'Cadillac only' medicine is nowhere more real than in the American hospital."²¹ Part of the reason for Cadillac only medicine may be that it tends to maximize the skilled labor input, and the physician's preferences may affect resource allocation. Nevertheless, a bias toward high quality hospitals is what the model would predict.²²

Third, evidence from the nursing home industry agrees with the prediction of the model that nonprofit institutions will emphasize quality. More nonprofit nursing homes have a registered nurse as the top nursing skill level than proprietary homes.²³ This is shown in Table 1.

¹⁹ See Somers and Somers (p. 77).

²⁰ Recall Long's observation, quoted earlier in this section, that "... hospitals seldom cut corners when doing so would reduce the quality of care."

²¹ See Gerald D. Rosenthal (p. 109).

²² Although nonprofit educational institutions are not all of high quality, this does not mean that the bias does not exist there also. Many of the observed quality differences in education may be due to differences in the position of the tradeoff curve; high endowment institutions have curves which are farther from the origin. The bias toward high quality which we believe exists for nonprofit institutions is for a given tradeoff curve. Since the variance in tradeoff curves for hospitals is probably less than the variance for educational institutions, the bias appears in the hospital sector as Cadillac only hospitals.

²³ See Jerry A. Solon. This index of quality is subject to the qualification noted above about indices of quality.

Last, the model predicts the often maligned duplication of sophisticated and expensive equipment in voluntary hospitals.²⁴ According to the de Bakey Commission on Heart Disease, Cancer, and Stroke, 30 percent of the 777 hospitals equipped to do open-heart surgery had no cases in the year under study. Of the 548 hospitals that did have cases, 87 percent did less than one operation per week.²⁵ The prestige accruing to the voluntary hospital possessing advanced equipment (and the value its decision maker gives that prestige) may also be the reason for greater capital intensity in the voluntary hospitals than the proprietaries.²⁶ An alternative explanation is that capital is relatively cheaper for the voluntaries so that cost minimization would lead to greater capital intensity.²⁷ Cost minimization, however, is neutral with respect to duplication of equipment, since output level is not specified. However, the interest of the decision maker in the voluntary hospital in quality as well as quantity may well be causing the duplication of equipment which in turn is causing the greater capital intensity.

In normative terms, the problem is that the decision-maker has picked a point on the quantity-quality tradeoff curve which is optimal for him but not necessarily socially optimal.²⁸ Are there any factors which might induce a socially optimal choice by the decision maker? First, the possibility of entry exists. We show below that this is unlikely to correct the problem. Second, philanthropy, insofar as its gifts are for a specific purpose, places a constraint on the point on the tradeoff

²⁴ See John T. Dunlop, and Rosenthal (p. 109).

²⁵ Cited in Rosenthal.

²⁶ Presumably more prestige comes from advanced capital equipment than additional labor. For figures on the greater capital intensity of the voluntaries, see Somers and Somers (p. 49).

²⁷ See Ralph E. Berry, Jr. (p. 117ff).

²⁸ The socially optimal point would be the outcome observed in a market dominated by knowledgeable consumers which functioned so as to satisfy their tastes (assuming the income distribution is optimal).

curve which the decision maker selects.²⁹ Unfortunately, the projects for which funds are given are usually those which will increase the quality of the output, whereas, as we have seen, the danger is that the decision maker will on his own choose too high a quality level. Thus, philanthropy does not help very much either in remedying this defect of the organization of the industry.³⁰

The second reason why nonprofit status hinders economic efficiency is the barriers to entry which result from it. Consider the possibility of entry by other hospitals. If highly unusual entry conditions existed, the hospital sector might still be considered efficient. Suppose that a hospital would enter if it thought it could undersell hospitals already in the industry, even though it would not make a profit from so doing, since it would set price at average cost. In other words, there would be a number of potential quantity-maximizing hospitals, all of whom now be outside the industry and so producing a zero quantity. Since they are quantity maximizers, they can better their condition by entering the industry and beginning production. In so doing, they force the price of hospitals already producing down to minimum average cost and force them to produce lower quality products. *Mirabile dictu*, the nonprofit hospital sector would be efficient.

While this type of entry may sound strange, before dismissing it out of hand, one should consider what attracts resources into the hospital sector (and other nonprofit sectors) now. Evidently it is not the opportunity for profit. Perhaps it is not too wide of the mark to suggest that the chance to provide a service will lead some

altruistic citizens to try to establish a hospital.³¹ This does not appear to be such a far fetched explanation for the founding of either hospitals or private colleges. They are often started by civic-minded individuals who organize a fund drive.

The problem is whether it is realistic to rely on this type of entry to assure satisfaction of demand and least cost production. For the possibility of entry by other nonprofit firms seems to be weak support indeed for the proposition that the hospital sector, left to its own devices, will reach a social optimum. It is one thing to say concern for the common weal is responsible for the establishment of hospitals, but quite another to say that it does so in a manner that we might term optimal.

We have spoken throughout in terms of a decision maker. Yet the theory of the firm assumes that an entrepreneur, not a decision maker or an administrator, will seize a profitable opportunity for entry. There is no position in the hospital (or nonprofit) sector for an entrepreneur. The entrepreneurs in this case are the civic-minded organizers, but they may not be present when an opportunity for entry exists. This has two consequences. Since there is no mechanism analogous to the possibility of profit which makes the threat of entry credible to existing institutions, there is less of a spur to least cost production. Inefficient but already established firms may continue to exist. Also, demand must generally be met by existing firms, which means some consumer demands for hospital care may not be met. If the nonprofit form of organization is to be retained, there is a need for government

²⁹ After receiving a tied philanthropic gift, the decision maker is in exactly the same position as a consumer who is given more of one particular good.

³⁰ However, if the philanthropist derives benefit from gifts which raise the quality level, optimality requires a higher quality level than if he did not. The argument here is that such gifts do not correct an already existing bias towards too high quality.

³¹ "Back in the nineteen-twenties, a group of eight or ten collectors in New York decided [to establish the Museum of Modern Art]. . . . The Metropolitan would not show any so-called modern art. They would not collect it and they would not show it, and one of the reasons the Museum of Modern Art was established was that this otherwise great institution was not ready to move into this field" Nelson A. Rockefeller.

policies to promote entry, for example, by continuing tax writeoffs to philanthropists or by providing funds for hospital construction.³²

However, there may be mechanisms other than entry by nonprofit hospitals for driving average revenue to minimum average cost and for satisfying consumer demands.³³ Suppose hospitals are being operated at an average cost above minimum average cost. An opportunity would exist for the entry of a profit-making firm. Why do we not observe this? That is, why does this sector of the economy continue to be organized primarily on a not-for-profit basis? One possibility is, of course, that hospitals are operating at minimum average cost and consumer demands are satisfied, so there is no opportunity for entry for a profit-making firm. If so, the existence of a profit-making sector keeps the nonprofit sector efficient in the same way that nonunion firms may place a constraint on wage changes in unionized firms. But this is at best an explanation of why nonprofit and profit-making firms might co-exist. It does not explain why so very few profit-making hospitals exist.

One possible explanation has been advanced by Kenneth Arrow. He views the notion of profit in medicine as incompatible with the delegation of choice the consumer makes to the physician; that is, as incompatible with the trust the consumer must have that the physician is deciding solely in the interests of the consumer. Hence, hospitals are nonprofit. This view, while not without some intuitive appeal, does not withstand close scrutiny. The argument about trust, if it is true at all, is only true for medicine. Yet schools and universities, museums, symphony orchestras, and theater groups are also often organized on a nonprofit basis. Hence,

³² Even if nonprofit status is not retained, the role physicians play in resource allocation may lead to a quality bias.

³³ Or at least to a point on the average cost curve if indivisibilities are important.

some explanation of the phenomenon more general than trust must be found. Also, the patient places his trust in the physician. It does not logically follow that the hospital or the medical school must be nonprofit. The existence of profit-making firms in medicine—most notably pharmaceutical manufacturers and distributors, but also private laboratories and private nursing homes—tends to indicate that there must be other reasons for hospitals' being nonprofit. Thus, we cannot rely on this hypothesis to explain why we do not observe the entry of profit making firms.

There are other, simpler explanations. One is merely that there may be legal barriers to entry for profit-making firms. Another has to do with the subsidies the hospital receives from private philanthropy and its favorable tax status.³⁴ This permits the hospital some deviation from the minimum average cost of a private firm and hence some scope for inefficiency. The nonprofit hospital might be run more inefficiently than a hypothetical profit-making one and yet be able to charge a price equal to or below that of a private firm. The difference in costs would simply be made up by the subsidy. Nor should it be thought that philanthropy is negligible; for all nonfederal short-term hospitals in 1964 there was a capital input of \$1.68 billion, 38 percent of which or over \$630 million came from philanthropy. The 38 percent is up from 25 percent in 1958.³⁵ The argument here is not that philanthropy directly causes inefficiency, but that it hinders selection of the fittest.³⁶ Good management is not always rewarded, which may explain the persistent calls for strengthening hospital management.³⁷

³⁴ This explanation is also cited by Arrow (p. 950).

³⁵ Irwin Wolkstein, cited in Somers and Somers (p. 211).

³⁶ There is some reason to think, however, that operation of the market may not lead to selection of the fittest even under competitive conditions in any simple way. See Sidney G. Winter, Jr.

³⁷ See Somers and Somers (pp. 51-55, 121-26, 247-50, 286-88, among others).

Philanthropy thus has its drawbacks. Perhaps philanthropy's favored tax status should be discontinued. Its existence, which provides some shelter for the non-profit hospital, raises the barriers to entry by profit-making firms.

The analysis up to this point has assumed that the consumer pays his own bill. The importance of third party schemes such as Medicare makes it important to modify the model to take account of their existence. It should be clear that so long as resources are constrained, the quality-quantity tradeoff is an inherent one; these are simply two commodities to which the decision maker can allocate his resources. The effect of changing the basis of payment may be merely to alter the location and shape of the tradeoff curve and doing that alters none of the conclusions reached above.

However, simple charge or cost reimbursement by a third party introduces a further potential inefficiency. Under this system the decision maker could conceivably push both quality and quantity to the point where the additional utility to him was zero. What would keep him from doing so? To do so would naturally lead to relatively high insurance rates, and this might lower the quantity demanded.³⁸ Would the decision maker take account of this effect? If there were many hospitals in one area, any one hospital's contribution to the high insurance rates would be negligible so that no single decision maker would take account of his contribution to them.³⁹ Thus, quality may be even higher and quantity lower than all would desire. If rates were based on the experience of a small number of hospitals, the effect each hospital has on insurance rates is more noticeable and so more likely to be taken

into account. The growth of hospital planning councils may be viewed as a step to make the large-group case similar to the small-group case.

One implication of the large-group case is that there is no incentive to the decision maker to minimize cost. In effect, resources are no longer constrained to the individual decision maker; the tradeoff curve can be shifted out at will. This, of course, is the extreme case: nevertheless, the real world may be quite close to it. Even in a world in which resources are treated as free, however, lower qualities may have a negative marginal benefit to the decision maker and so not be produced. Thus, the cost-reimbursement model can generate the same quality bias that the demand-curve model shows.

To summarize the conclusions of this discussion: We have examined ways in which the nonprofit status of voluntary hospitals may cause misallocation of resources. First, there is a bias against producing lower quality products (a bias in the sense that a profit-maximizing firm would produce such qualities). Second, there is little reason to think that a nonprofit hospital will enter in response to a profitable opportunity (which may exist either because consumer demands are not being satisfied or because inefficient hospitals are providing the product). Philanthropy gives the nonprofit hospital some latitude for inefficiency, and this, among other things, tends to forestall entry by profit-making firms. An additional problem exists if the hospital is simply reimbursed by a third party for its costs. By removing the budget constraint, incentives for least cost production are weakened.

REFERENCES

- K. J. Arrow, "Uncertainty and the Welfare Economics of Medical Care," *Amer. Econ. Rev.*, Dec. 1963, 53, 941-73.
 W. J. Baumol and W. G. Bowen, "On the

³⁸ It might also have the opposite effect. By making insurance more attractive to previously uninsured risk-aversers, effective price could be lowered and demand raised.

³⁹ See Mancur Olson, Jr. for an analysis of this phenomenon.

- Performing Arts," *Amer. Econ. Rev. Proc.*, May 1965, 55, 495-502.
- R. E. Berry, Jr., "Competition and Efficiency in the Market for Hospital Services," unpublished doctoral dissertation, Harvard Univ. 1965.
- R. E. Brown, "Let the Public Control Utilization Through Planning," *Hospitals*, Dec. 1959, 33, 34-39.
- J. A. Crabtree, "Plans for Tomorrow's Needs in Local Public Health Administration," *Amer. J. Public Health*, Aug. 1963, 53, 1175-81.
- J. S. Davis, *Essays on the Earlier History of American Corporations*, Cambridge, Mass. 1917.
- J. P. Dixon, "Hospitals and the Community," Eugene A. Confrey, ed. *Administration of Community Health Services*, Chicago. International City Managers' Association 1961, 262-84.
- J. T. Dunlop, "The Capacity of the United States to Provide and Finance Expanding Health Services," *Bull. N. Y. Acad. Medicine*, Dec. 1965, 41, 1325-37.
- M. S. Feldstein, *Economic Analysis for Health Service Efficiency*, Amsterdam 1967.
- M. Gilbert, "The Problem of Quality Changes and Index Numbers," *Econ. Develop. Cult. Change*, Apr. 1961, 9, 287-94.
- J. Hirschleifer, "The Exchange Between Quantity and Quality," Santa Monica, The RAND Corporation (P-406) 1955.
- M. F. Long, "Efficient Use of Hospitals," S. J. Axelrod, ed. in *The Economics of Health and Medical Care*, Ann Arbor 1964, 211-26.
- F. Machlup, "Theories of the Firm: Marginalist, Behavioral, Managerial," *Amer. Econ. Rev.*, Mar. 1967, 57, 1-33.
- J. L. Nicholson, "The Measurement of Quality Changes," *Econ. J.*, Sept. 1967, 77, 512-30.
- M. Olson, Jr., *The Logic of Collective Action*, Cambridge, Mass. 1965.
- M. V. Pauly, "Efficiency in Public Provision of Medical Care," unpublished doctoral dissertation, Univ. Virginia 1967.
- N. A. Rockefeller, "The Governor Lectures on Art," *New York Times Magazine*, Apr. 9, 1967.
- G. D. Rosenthal, "The Public Pays the Bill," *Atlantic Monthly*, July 1966, 218, 107-10.
- J. Rothenberg, "Welfare Implications of Alternative Methods of Financing Medical Care," *Amer. Econ. Rev. Proc.*, May 1951, 41, 676-87.
- J. A. Solon, "Nursing Homes and Medical Care," *New England J. Medicine*, Nov. 14, 1963, 269, 1067-74.
- H. M. Somers and A. R. Somers, *Medicare and the Hospitals*, The Brookings Institution, Washington 1967.
- C. Stevens, "Hospital Market Efficiency: The Anatomy of the Supply Response," paper presented at the second Conference on the Economics of Health, Baltimore, Dec. 5-7, 1968.
- H. Theil, "Qualities, Prices, and Budget Inquiries," *Rev. Econ. Stud.*, 1952, 19, 129-47.
- S. G. Winter, Jr., "Economic Natural Selection and the Theory of the Firm," Santa Monica, The RAND Corporation (P-2167) 1960.
- U.S. Department of Labor, Bureau of Labor Statistics, *Employment and Earnings for the United States, 1909-1962*, Washington 1963.

Cyclical Variation in the Interindustry Wage Structure

By MICHAEL L. WACHTER*

This paper presents an analysis of cyclical variation in the interindustry wage structure. The study is therefore concerned with changes in relative wages among industries and not directly with the determination of the absolute wage level or changes in the absolute wage level. The theory of the wage structure presented here, is based on two hypotheses. The first, introduced initially as a maintained hypothesis, is that differences among industries in the competitive structure of their product and labor markets result in wage inequalities for given skill levels. The second hypothesis, and the major finding of the paper, is that the size of the wage differentials is a function, not only of the extent and nature of the noncompetitive elements, but also of the excess demand conditions in the labor market. This result is derived from the proposition that under competitive supply conditions, labor allocation is based on relative wages.

To support the theory, a wage structure equation is estimated with the coefficient of variation of straight time, average hourly earnings in the two-digit manufacturing industries as the dependent variable.¹ The implications of the study for

aggregate and industry wage equations are then discussed.

I. *The Theory of the Interindustry Wage Structure*

The theory of the interindustry wage structure presented in this paper, rests on two hypotheses. The first is that product and labor markets among industries result in wage inequalities for given skill levels.

The accuracy of this hypothesis is a topic of debate in the literature.² It is clear, from simple inspection of wage rates, that differentials exist and that these differentials have been present for too long to be considered a random disturbance from the competitive ideal. The claim, however, has been made that these differentials simply reflect differences in productivity among workers. Most empirical studies, on the other hand, find that both elements are present: that high-wage industries obtain more productive workers, but that the differential cannot be explained solely by these factors.³

The two noncompetitive elements that are particularly stressed in the literature are "ability to pay" considerations (emanating from the structure of the product market) and the strength and degree of unionization. No attempt is made in this paper to assess the relative importance of the noncompetitive elements. Since there

* The author is an assistant professor at the University of Pennsylvania. He is indebted to John T. Dunlop, Otto Eckstein, Peter Doeringer, and Lester Thurow for helpful suggestions. He also thanks the editor for correcting an error in an earlier version of the paper. The work was done under a Labor Department dissertation grant.

¹ The coefficient of variation is obtained by taking, for each year, the simple average of the two-digit earnings series in manufacturing and then computing the standard deviation. The coefficient of variation is then the ratio of the standard deviation over the mean. The original earnings series are adjusted for overtime.

Since no overtime earnings series is available for printing, the industry is omitted.

² See, for example, Adrian Throop (1968), Leonard Weiss, and Frank Stafford.

³ Weiss, for example, finds that high-wage industries hire "superior" workers, but that many of the traits that define superiority, do not lead to greater productivity.

are high intercorrelations for manufacturing industries among such variables as degree of unionization⁴ and concentration ratios in the product market, I simply assume that it is some combination of non-competitive product and labor market influences that affects the wage premium.⁴

The noncompetitive elements are relatively stable over the short run and have shown little variation over the postwar period. This suggests that these factors explain the existence of wage premiums, but not changes in the premiums over the postwar period. That is:

$$(1) \quad CV_t = a_0$$

where CV , the coefficient of variation described in fn. 1, is a statistic of the dispersion of the wage structure and a_0 is the (constant) size of the inequalities resulting from the factors discussed above.

The second hypothesis is that the size of the wage inequalities in the interindustry wage structure, outside of this constant, are largely a function of the excess demand conditions in the labor market. This can be expressed:

$$(2) \quad CV_t = a_0 + a_1 U_t$$

where U , the aggregate unemployment rate, represents the excess demand conditions in the labor market.⁵

The rationale for equation (2) is found by analyzing the supply curve of labor to the various industries. In an economy where competitive forces do not equalize wages (for equal skill) and aggregate un-

employment exists, the supply function can be specified in the following form:

$$(3) \quad SS_{i,t} = f(W_{i,t}/W_t^*; U_t)$$

where SS_i is the supply to the i th industry, W_i is the wage paid by the i th industry, and W^* is a proxy variable representing wage rates existing elsewhere in the system.⁶ Taking the variables to be in logarithmic form, the supply function is assumed to have the following shape:

$$(4a) \quad \frac{\partial SS_i}{\partial (W_i/W^*)} > 0$$

$$(4b) \quad \frac{\partial SS_i}{\partial U} > 0$$

$$(4c) \quad \frac{\partial^2 SS_i}{\partial (W_i/W^*)^2} > 0$$

$$(4d) \quad \frac{\partial^2 SS_i}{\partial (W_i/W^*) \partial U} < 0$$

Inequalities (4a) and (4c) state that supply is positively related to the relative wage and that the higher the relative wage, the greater is the elasticity of the supply curve. Inequality (4b) says that an increase in unemployment increases the supply of labor available to an industry. The final inequality states that the higher the relative wage, the smaller the effect of a change in the unemployment rate. The fact that the supply of labor is not infinitely elastic to every industry, at a given positive unemployment rate, is due to the traditional causes of short-run inelasticity, including the observed downward rigidity of the supply price of unem-

⁴ John Dunlop's "wage contour" mechanism provides the underlying structure for this hypothesis.

⁵ In a sense, a test of equation (2) is also a test of the first hypothesis that wage premiums exist for equal skill. If the observed wage differences simply reflect quality differences, the coefficient of variation should not exhibit cyclical variation. That is, wage inequalities should not vary systematically with the unemployment rate, because the supposedly low-wage industries will be as content to hire low-quality workers in tight as in loose labor markets.

⁶ The supply function, equation (3) needs to be modified to account for the influence of the "internal labor market" of the firm. Due to specific training and the industrial practices of hiring at the bottom and promoting from within, a worker may be able to earn a higher wage in a low-wage industry, where he is presently employed, than he would elsewhere. The result is that a certain percentage of an industry's employment is "locked-in." This obviates the necessity of introducing a size (of the industry) variable into the supply curve.

ployed labor⁷ (see Hirshel Kasper and Robert Lucas and Leonard Rapping).

The relative wage of a particular industry then determines the supply available to that industry at the prevailing unemployment rate. This does not mean that the available supply is, in fact, employed by the industry in question. By definition, whenever an industry pays a wage premium, its relative wage is above the wage required to attract the number of workers actually employed. That is, high-wage industries can be viewed as hiring off a labor queue. It is probable that almost all industries pay a premium in order to reduce uncertainties attached to unexpected fluctuations in their demand or supply of labor. For purposes of exposition, however, it is assumed that only high-wage industries pay a wage premium.

Cyclical variation in relative wages can be analyzed in terms of shifts in the industry supply curves due to changes in aggregate unemployment. As labor markets tighten, for example, low-wage industries find a decline in the size of the labor force available to them. This involves not only an increasing difficulty in attracting new workers, but also a greater problem in retaining workers already hired. The result is that low-wage industries attempt to improve their competitive standing by narrowing the wage structure. (Note that the increase in the desired relative wage does not depend upon an increase in the demand for labor by the particular industry. If such an increase occurs, the desired increase in the relative wage is that much

larger.) Similarly, as unemployment rises, low-wage industries find it increasingly easy to maintain a given labor supply, so that they attempt to reduce their relative wage. Whether or not the competitive industries are successful in this cyclical relative wage change pattern depends upon the action of the high-wage industries.⁸

It is clear that high-wage industries, as well as low-wage industries, experience cyclical variability in their labor supply. The former, however, pays a wage premium and thus has some discretion in the timing and amount of its wage changes. A discretionary policy, however, need not be acyclical. That is, high-wage industries may attempt to at least maintain their relative wage position over the cycle.

There are, however, a number of advantages to the high-wage industries in adopting a wage change policy that permits them to maintain their ranking in the wage structure, but also allows their percentage differentials to vary inversely with the cycle. The traditional explanations are related to noncompetitive elements in the labor market. First, trade unions appear to be more interested in wage developments within the high-wage group than outside that group.⁹ Second, unions are interested

⁸ This theory has certain similarities with Melvin Reder's theory of the determination of occupational wage differentials. His theory predicts that wage differentials narrow during expansions and widen during recessions. The cause of this behavior is the ability of skilled workers to fill less skilled jobs during recession and the necessity of upgrading lower skilled workers during expansions. Since, in manufacturing (the source of data for this study) the high-skilled industries are also the high-wage industries, the two theories complement each other.

The amount of cyclical variation among manufacturing industries, explained solely by changes in occupational wages, however, may not have been significant during much of the postwar period. See George Hildebrand and George Delehanty.

⁹ This notion has been formalized in somewhat different ways by Dunlop in the wage contour mechanism and by Otto Eckstein and Thomas Wilson in the "key group" bargaining.

⁷ An explanation is needed as to why a relative demand variable is not included in the analysis. The reason is inherent in the specification of the supply curve. Due to the payment of wage premiums, changes in demand have a relatively minor impact on the desired relative wage of the high-wage industries. Therefore a demand variable for the low-wage industries may be more appropriate than a relative demand variable. The unemployment rate, however, may serve as a proxy for this effect.

in a certain minimum rate of wage increase during loose labor markets, even at the expense of a somewhat slower rate of growth during tighter labor markets; see George de Menil and Throop (1968). Finally, contract bargaining introduces a lag in the response of the high-wage industries. One impact of the lag is to smooth cyclical fluctuations.¹⁰

An additional explanation may be based on the noncompetitive elements in the product market. An acyclical pattern may be a natural complement to target return pricing. Eckstein and Fromm, among others, have shown that there is a high coincidence of noncompetitive elements in the product market and the use of target pricing rather than competitive pricing. Target pricing has two elements that are relevant to this analysis: product prices do not change frequently; and product prices are geared to standard labor costs and not to changes in product demand. This suggests that for an optimal wage policy, the wage component of labor costs should be, as far as possible, predictable into the future and insensitive to short-run cyclical changes in demand. If labor costs are sensitive to excess demand conditions in the labor market, then target return pricing loses its distinguishing features.

To summarize, because of factors in the product and/or labor markets of high-wage industries, wage changes in these industries are likely to be relatively acyclical. The result is that low-wage industries successfully adjust their relative wage position in response to cyclical variation in a competitive labor supply framework.¹¹

¹⁰ The major work on the impact of unions on relative wages is the study by H. Gregg Lewis. His book contains other references.

¹¹ One major factor, not dealt with in this paper, is that low-wage industries may react to a tightening labor market by decreasing quality rather than by increasing wages. Most low-wage industries probably utilize both methods of adjustment. The industries that rely heavily on the quality adjustment, however, may be the local service industries.

II. *Monetary Factors and Distributed Lag Elements in the Wage Structure*

Equation (2) specifies that the wage structure is determined solely by real factors. This simply means that relative prices (of labor) are not a function of the absolute price level.¹² A cost of living variable, however, may affect the wage structure in disequilibrium. That is, industries may differ in the speed with which they adjust wages to changes in the cost of living, but not in the overall size of the adjustment. These differences are based on institutional arrangements in the product and labor markets. A great deal has been written on these factors, yielding a myriad of explanations as to how changes in consumer prices affect the wage structure. Martin Bronfenbrenner and Franklyn Holzman in their review article on inflation, suggest that high-wage industries respond more rapidly during slow inflation and less rapidly during rapid inflation, than low-wage industries.¹³ This hypothesis can be tested by the addition of a nonlinear price change variable to equation (2).

In the preceding discussion, little explicit mention has been made of the speed with which the coefficient of variation adjusts to the unemployment rate and to changes in the cost of living. A relatively slow adjustment path is to be expected however for three major reasons. First, there are institutional lags, such as the dating of contracts and the existence of wage rounds.¹⁴ Second, there are the eco-

¹² A number of recent industry wage models do not have this property. This occurs in equations where the rate of wage change of an industry (or group of industries) is a function of the rate of change of consumer prices. Since the estimated coefficients for the price term generally differ among industries, such a model in effect, explains relative wages with absolute prices.

¹³ Their explanation is based on factors such as contract lags.

¹⁴ Eckstein and Wilson suggest that the real observation for wage changes is the wage round. Their longest wage round is four years. If their hypothesis is correct, annual equations should find relatively long lags.

nomic costs associated with large and frequent changes in relative wages. It must be remembered that industries need to adjust not only their wage level, but also their relative wage. Immediate adjustment would require, for example, no money wage rigidity on the downside and perfect knowledge of intended wage changes in other industries. Third, as previously mentioned, there are short-run inelasticities in the adjustment of the labor supply.

III. *Comparison with the Literature*

The determination of changes in the interindustry wage structure has previously been studied in some detail. The purpose of this section is not to summarize the extensive literature, but to present a sample of those studies that are most closely associated with the ideas presented in this paper.¹⁵

A widely accepted theory is that changes in relative wages result from differential rates of employment growth among industries. In a competitive labor market where all industries pay equal wages for equal skill, this theory states that short-run inelasticities in the supply of labor cause industries with the highest growth rates of employment to increase their relative wage (at least in the short run) to attract the necessary labor.

The above hypothesis has recently been tested by an *OECD* study group using data for the United States and for a number of other western economies.¹⁶ The authors of

the study found that there is "... no evidence of a strong systematic relationship between changes in earnings among individual industries and variations in relative employment" (p. 16). Their explanation for this result is that the role of allocating labor is performed by differentials in the existing level of wages. From this, they conclude that high wage industries may expand (relative) employment without increasing their relative wage. For low-wage industries, however, "... above average increases in wages could well be a condition for an expansion of numbers employed" (p. 19).

The *OECD* hypothesis differs from the theory presented here, in that this paper emphasizes shifts of the supply curve of labor, resulting from changes in excess demand conditions in the economy as a whole, rather than shifts of the demand curve of a particular industry. That is, the wage dispersion variable is a function of the total unemployment rate (and of changes in consumer prices) and not of the presence and location of differential changes in employment.¹⁷

The point that the relative wage may increase in the low-wage sector during tight labor markets has been noticed previously by other authors. Miss Haddy and Tolles, for example, found that the wage structure narrowed during the 1940's and ascribed this development to the tight labor markets of the period. More recently, the Council of Economic Advisers has suggested that with today's tight labor markets "... many workers at the bottom of the economy's wage structure now face opportunities of moving into more ad-

¹⁵ See, for example, Doris Eisemann, Pamela Haddy and N. Arnold Tolles, Lewis, Throop (1968), Lloyd Ulman, and *OECD*. The notion that changes in the wage structure are primarily due to the behavior of unions is reviewed in Lewis and will not be discussed here in detail. Lewis finds first, that unions secure a positive differential for their members over nonmembers, and second, that this differential narrows during expansions. He also claims that it is the dynamics of union behavior that causes the cyclical change in the differential. Although the theory presented here is not incompatible with Lewis' results, the former stresses the behavior of the low wage, nonunion sector in determining changes in the wage structure.

¹⁶ See Ulman, for further tests, using data for the

United States. Ulman is also one of the authors of the *OECD* report.

¹⁷ A second difference between this paper and the *OECD* report, as well as many of the earlier studies, is that the latter primarily utilize correlation analysis as a statistical tool. This makes it difficult to test for the existence of distributed lags. As will be seen in the next section, these lags are important.

TABLE 1—ESTIMATE OF EQUATION (6), ANNUAL DATA, 1947-67*

Lag	U^{-1}		\dot{P}		\dot{P}^2	
	Coefficient	T-value	Coefficient	T-value	Coefficient	T-value
0	-.8743	9.25	.1547	2.28	-1.6241	2.59
1	-.6779	14.27	.1768	3.45	-2.2371	5.60
2	-.4668	8.57	.1724	3.54	-2.3812	6.58
3	-.2408	5.34	.1414	3.21	-2.0564	5.98
4			.0841	2.92	-1.2626	5.39
Sum	-2.2599		.7293		-9.5614	
$\bar{R}^2 = .9889$		Standard Error = .002259		DW = 2.341		
Constant term = .2132		Korean War Dummy = -.01317		(6.66)		
(54.16)						

* Coefficient of Variation includes two-digit manufacturing industries, except Printing. Sources: Average hourly earnings from 1947-55 and average hourly earnings, excluding overtime, from 1956-67 are contained in the *Employment and Earnings Statistics for the United States, 1909-1968*. The method of converting the 1947-55 data to straight time earnings is described in the *Monthly Labor Review*, May 1950. For the unemployment rate, see *Employment and Earnings* and for the percentage change in the consumer price index, see the *President's Economic Report*.

vantageous jobs. Accordingly, wage increases in the low paid sectors are likely to exceed the average wage rise in the economy as a whole" (1966, p. 79).

IV. Empirical Results

The above discussion suggests the following equation:

$$(5) \quad CV_t = a_0 + \sum b_i U_{t-i} + \sum c_i P_{t-i} + \sum d_i \dot{P}_{t-i}^2$$

where the summation is over years. The variable \dot{P} is the percentage change in the consumer price index.

In arriving at a final equation, two final additions are made. First, the reciprocal of the unemployment rate is substituted for the unemployment rate. The rationale for this change is the assumption that at some high level of unemployment, wage dispersion becomes relatively insensitive to further increases in unemployment. This may be due, for example, to a decline in market or union power during periods of high unemployment. (The empirical results are not sensitive to the particular specification of the unemployment variable.) Secondly, a Korean War dummy

variable is added. The equation to be estimated is then:

$$(6) \quad CV_t = a_0 + a_1 KW_t + \sum b_i U_{t-i}^{-1} + \sum c_i P_{t-i} + \sum d_i \dot{P}_{t-i}^2$$

Equation (6) is fitted to annual data for the period 1947-67. The distributed lags are estimated using the Almon technique.¹⁸ The results are presented in Table 1.¹⁹ As can be seen, the equation describes the data with an adjusted \bar{R}^2 of .9889. The unemployment variable is highly significant, providing empirical justification for the labor supply mechanism presented in the initial section. The long-run elasticity at the means is approximately 0.25. The

¹⁸ See Shirley Almon for an explanation of the technique. For each independent variable estimated with a lag, two Almon variables are used. This permits the estimation of a second degree polynomial with a zero value assumed for the last lag weight. The length of the lags do not interfere with the 1947 starting date for the regression, since values of the independent variables are available for the period preceding 1947.

¹⁹ Since the coefficient of variation is a slow moving variable, it is possible that its movements can be explained by a trend variable as well as by the actual independent variables appearing in Table 1. To test this, a linear trend is added to equation (6). The T -statistic is below unity and the adjusted \bar{R}^2 falls slightly. Similar results are obtained for a log-linear trend.

length of the unemployment rate lag to 50 percent adjustment is between one and two years.

The price terms are also significant and imply the nonlinear response predicted by Bronfenbrenner and Holzman. The coefficients for the price terms suggest the following conclusions. First, inflation has a net widening effect on the wage structure except at very rapid rates of inflation. The maximum positive effect, however, occurs below a 4 percent inflation rate. After that point, increases in the rate of price change result in decreases in the coefficient of variation. Second, the size of the overall price effect on the wage structure is quite small. The elasticity of the coefficient of variation, measured at its mean, with respect to the total price effect, measured at its point of maximum positive effect is only .09.²⁰

V. *The Relationship Between Earnings and Wages for Equal Work*

The theory of the interindustry wage structure, presented in this paper, states that wage premiums (for equal skill) vary over the cycle. In the test of the theory, straight time average hourly earnings are used as the proxy for wages paid for equal skill. Differences in average hourly earnings among industries, however, are primarily a function of differences in the skill mix among industries (see Throop 1968). Thus, it is at least possible that an indus-

try that has relatively high average hourly earnings may have an even higher relative skill level, so that the industry is below average in wages paid for equal skill. Accordingly, this industry should be considered a low-wage industry and the use of average hourly earnings in place of wage rates paid for equal skill, in this case, creates a serious bias. The purpose of this section is to show that the bias, to the extent that it does exist, is not serious.

Evidence that the level of average hourly earnings in a cross section of two-digit industries is correlated with the premium paid for equal work can be obtained from a regression of average hourly earnings on the skill mix variable constructed by Throop (1968). Throop's equation for the manufacturing sector alone is presented in Table 2.

TABLE 2—*ln* OF AVERAGE HOURLY EARNINGS, *W*, REGRESSED ON *ln* OF INDEX OF SKILL LEVEL, *S*

1950	$\ln W = -8.3063 + 1.1137 \ln S$	$\bar{R}^2 = .6470$ (5.67)
1960	$\ln W = -10.7085 + 1.3892 \ln S$	$\bar{R}^2 = .7156$ (6.62)

Note: *T*-statistics are in the parentheses.

Source: The skill data were kindly provided by Throop.

As can be seen, the coefficient of the skill term is greater than unity in both equations. This indicates that the higher the skill level, the greater the wages paid for a given skill mix (see George Borts). The coefficient for the 1950 equation, however, is slightly less than one standard error from unity.

An inspection of the residuals of the 1950 equation, however, shows that only one industry, apparel, is largely responsible for the low value of the slope. The skill variable in the 1950 equation, with apparel omitted, has a coefficient of 1.49 with a standard error of .20.

This fact suggests that the coefficient of variation be recalculated, omitting

²⁰ Caution, however, must be exercised in interpreting the quadratic price variable. The problem is that the price change variable does not take on a value between 3.4 percent and 7.7 percent during the estimation period, 1947-67. Furthermore, a majority of the observations are below 2.5 percent. The quadratic form is thus estimating a positive relationship between inflation and interindustry wage dispersion at very low rates and a negative relationship at very high rates. In between these extremes, the predicted relationship is due simply to the specific functional form of the variable. As a result, it may be difficult to distinguish a quadratic form from an alternative function in \hat{P} that is negatively sloped throughout and concave to the origin. For example, if the linear price variable in equation (6) is omitted, the adjusted \bar{R}^2 only declines to .9808.

TABLE 3—ESTIMATE OF EQUATION (6), ANNUAL DATA, 1947-67*

Lag	U^{-1}		P		P^2	
	Coefficient	T-value	Coefficient	T-value	Coefficient	T-value
0	— .7413	7.38	.1488	2.06	—1.4687	2.20
1	— .5251	10.40	.1530	2.81	—1.7877	4.21
2	— .3294	5.69	.1403	2.71	—1.8003	4.68
3	— .1544	3.22	.1105	2.36	—0.9065	3.64
Sum	—1.7502		.6163		—7.4699	

$\bar{R}^2 = .9772$	Standard Error = .002401	$DW = 1.729$
Constant term = .2008	Korean War Dummy = —.001132	
(47.98)	(5.38)	

* Coefficient of Variation includes two-digit manufacturing industries, except Printing, Apparel, Ordnance, Instruments, and Miscellaneous Manufactures.

Sources: See Table 1.

apparel and the three industries for which skill data are not available. For the remaining industries, the results presented above provide support for the use of straight time average hourly earnings as a proxy for wage rates for equal skill. The results obtained by fitting equation (6) to the revised dependent variable, are shown in Table 3. Although the adjusted \bar{R}^2 is reduced somewhat from its level in Table 1, the independent variables remain highly significant with the correct sign. The results, therefore, largely confirm the findings of Table 1.

VI. Summary and Implications

This paper has attempted to show that the interindustry wage structure, represented by the coefficient of variation among two-digit industries, is a distributed lag function of the unemployment rate and changes in the cost of living. The unemployment rate operates as a proxy for the excess demand conditions existing in the labor market. During tight labor markets, for example, low-wage industries are forced to narrow the wage differential between themselves and the high-wage industries in order to attract and hold the desired labor supply. The changes in the cost of living variable, on the other hand, enters the equation because industries may differ in

the speed with which they respond to inflation. The results indicate that except for very rapid inflation, high-wage industries respond more quickly to price changes than do the low-wage industries.

One implication of this study is that monetary and fiscal policy, aimed at choosing a point on the hypothesized unemployment-inflation trade off, has an effect on the income distribution among wage earners as well as between wage earners and the unemployed. Specifically, emphasis on price control, rather than stress on maintaining low rates of unemployment, results in a widening of the wage structure, benefiting those in the high skilled, high paying industries relative to those at the bottom of the wage structure. On the other hand, programs designed to achieve price stability and low unemployment should have a leveling effect on interindustry wages. This results from the much greater elasticity of the wage structure to unemployment than to price changes. Table 4 contains predicted levels of the coefficient of variation for different target values of the unemployment rate and the percentage change in the cost of living.²¹

²¹ The tables are derived from the coefficients reported in Table 1. It is assumed that the independent variables maintain the particular values for at least as long as the lag length.

TABLE 4—PREDICTED VALUES OF THE COEFFICIENT OF VARIATION
FOR TARGET VALUES OF THE INDEPENDENT VARIABLES

	Unemployment Rate	Annual Change CPI	Predicted Value CV
1	8 Percent	2 Percent	.1940
2	5	3	.1813
3	3	3	.1542
4	3	5	.1540
5	2	8	.1074

T. Paul Schultz, in his study of cyclical changes in the income distribution of the Netherlands economy, obtained results similar to those reported here. In particular, he found that low unemployment reduced the income inequality and that price changes had only a minor effect.

A second implication of this study concerns the recent debate over the effectiveness of the guideposts policy during the 1960's (see Paul Anderson, Perry (1967, 1968), Throop (1969), and Wachter). It has been suggested that one indication of the success of the guideposts policy was that the rate of wage change in the visible industries (those industries that are likely to be most affected by the guideposts) was slowed, relative to the wage change in the invisible industries. The visible industries, however, are in general the high-wage group and the invisible industries are substantially the low-wage group (see Wachter.) As a result, the guideposts hypothesis and the theory suggested in this paper coincide in predicting a narrowing of the wage structure in the 1960's.²²

A third implication of this study concerns the specification of aggregate wage change models. This study suggests that the amount of wage dispersion may be a significant variable in the determination of the rate of aggregate wage change.²³ The

theory presented in Section I states that low-wage industries attempt to alter their relative wage position to meet changes in the tightness of the labor market. They succeed in their goal because of the relatively acyclical wage change policy of the high-wage industries. The possibility is raised, however, that as the differentials narrow, the high-wage industries may attempt to prevent a further narrowing in order to protect their preferential position in the labor market. This would lead to an increase in the rate of aggregate wage change.

This, for example, may help to account for the relatively slow rate of wage change in the tight market of the 1964-67 period compared with the tight market of the late 1940's and early 1950's. The reason is that the wage dispersion as measured by the coefficient of variation was larger in the 1960's than in the earlier period.²⁴

If the unemployment rate remains below 4 percent, however, the coefficient of variation (operating with a distributed lag adjustment to the unemployment rate) should continue to fall. To the extent that the high-wage industries attempt to protect their position in the interindustry wage structure, the decline in the wage dispersion will lead to an increase in the rate of aggregate wage change.

²² A guideposts dummy variable was added to equation (6), but proved insignificant.

²³ To my knowledge, no wage equation presently includes a dispersion index as an independent variable. The point has been previously discussed however, by Throop (1968).

²⁴ To provide an initial test of this hypothesis, the coefficient of variation, in four quarter moving average form is added to Perry's aggregate wage change model, see Perry (1966). The variable is significant with the correct negative sign at the 99 percent level.

REFERENCES

- S. Almon, "The Distributed Lag Between Capital Appropriations and Expenditures," *Econometrica*, Jan. 1965, 33, 178-96.
- P. T. Anderson, "Wages and the Guideposts: Comment," *Amer. Econ. Rev.*, June 1969, 59, 351-54.
- G. H. Borts, "Analysis of Interstate Income Differentials: Theory and Practice: Comment," *Regional Income*, Nat. Bur. Econ. Res. Stud. in *Income and Wealth*, Vol. 21, Princeton 1957.
- M. Bronfenbrenner and F. Holzman, "A Survey of Inflation Theory," in *Surveys of Economic Theory*, Vol. 1, New York 1965.
- G. de Menil, "Nonlinearity in a Wage Equation for the United States," *Rev. Econ. Statist.*, forthcoming.
- J. T. Dunlop, "The Task of Contemporary Wage Theory," in G. Taylor and F. Person, eds., *New Concepts in Wage Determination*, New York 1957.
- O. Eckstein and G. Fromm, "The Price Equation," *Amer. Econ. Rev.*, Dec. 1968, 58, 1158-83.
- O. Eckstein and T. A. Wilson, "Determination of Money Wages in American Industry," *Quart. J. Econ.*, Aug. 1962, 76, 379-414.
- D. Eisemann, "Inter-Industry Wage Changes, 1939-1947," *Rev. Econ. Statist.*, Nov. 1956, 38, 445-48.
- P. Haddy and N. A. Tolles, "British and American Changes in Interindustry Wage Structure under Full Employment," *Rev. Econ. Statist.*, Nov. 1957, 39, 408-14.
- G. H. Hildebrand and G. E. Delehanty, "Wage Levels and Differentials," in R. A. Gordon and M. S. Gordon, eds., *Prosperity and Unemployment*, New York 1966.
- H. Kasper, "The Asking Price of Labor and the Duration of Unemployment," *Rev. Econ. Statist.*, May 1967, 49, 165-72.
- H. G. Lewis, *Unionism and Relative Wages in the United States*. Chicago 1963.
- R. E. Lucas and L. A. Rapping, "Real Wages, Employment, and Inflation," mimeo.
- G. Perry, *Unemployment, Money Wage Rates, and Inflation*. Cambridge 1966.
- , "Wages and the Guideposts," *Amer. Econ. Rev.*, Sept. 1967, 57, 897-904.
- , "Wages and the Guideposts: Reply," *Amer. Econ. Rev.*, June 1969, 59, 365-70.
- G. Pierson, "Union Strength and the U.S. 'Phillips Curve'," *Amer. Econ. Rev.*, June 1968, 58, 456-67.
- M. W. Reder, "The Theory of Occupational Wage Differentials," *Amer. Econ. Rev.*, Dec. 1955, 45, 833-52.
- T. P. Schultz, "Secular Equalization and Cyclical Behavior of Income Distribution," *Rev. Econ. Statist.*, May 1968, 50, 259-67.
- N. J. Simler and A. Tella, "Labor Reserves and the Phillips Curve," *Rev. Econ. Statist.*, Feb. 1968, 50, 32-49.
- F. P. Stafford, "Concentration and Labor Earnings: Comment," *Amer. Econ. Rev.*, Mar. 1968, 58, 174-80.
- A. W. Throop, "The Union-Nonunion Wage Differential and Cost-Push Inflation," *Amer. Econ. Rev.*, Mar. 1968, 58, 79-99.
- , "Wages and the Guideposts: Comment," *Amer. Econ. Rev.*, June 1969, 59, 358-65.
- L. Ulman, "Labor Mobility and the Industrial Wage Structure in the Postwar United States," *Quart. J. Econ.*, Feb. 1965, 79, 73-97.
- M. L. Wachter, "Wages and the Guideposts: Comment," *Amer. Econ. Rev.*, June 1969, 59, 354-58.
- L. Weiss, "Concentration and Labor Earnings," *Amer. Econ. Rev.*, Mar. 1966, 56, 96-117.
- Organization for European Cooperation and Development, *Wages and Labour Mobility*. Paris 1965.
- U.S. Council of Economic Advisers, *The Annual Report*, Washington 1966.
- , Washington 1968.
- U.S. Department of Labor, Bureau of Labor Statistics, *Employment and Earnings Statistics for the United States, 1909-1968*. Bull. No. 1312-6. Washington 1968.
- , *Employment and Earnings*, Mar. 1968, 14.
- , *Mon. Lab. Rev.*, May 1950, 73, 537-540.

Monetary Growth Theory in Perspective

By JEROME L. STEIN*

"The rivalry of scholars increases wisdom . . ." *Babylonian Talmud*

Monetary growth theory is concerned with the role of money in a growing economy. Money is a medium of exchange and store of value which may or may not be costless to produce, and which is a liability of either the government or a private banking system. Monetary policy is concerned with the management of these types of money. Several questions immediately arise. To what extent can financial policies and institutional arrangements affect the time profiles of the capital-labor ratio $k(t)$, the real wage $w(t)$, and the rent per unit of capital $r(t)$? It is obvious that the growth of a commodity money (e.g., gold) will affect the real variables in the system, because resources (labor, capital) are required for the production of gold. Can variations in the rate of growth of a type of money which is costless to produce affect the time profiles, and steady state solutions, of these real variables? Is there an optimum growth of the various types of money? What are the most desirable stabilization policies in a growing economy?¹ These are some of the central issues in monetary growth theory.

There are several different ways of analyzing the effects of monetary policy in a growing economy. First, why restrict the analysis to a money which is costless to produce? Jürg Niehans considered the case where some fraction of the stock of money

consists of monetary gold. The opportunity cost of producing gold (or exports to purchase gold) is output which is no longer available for consumption and investment. The realism of this assumption is unquestionable and paves the way for a discussion of growth in an open economy which uses money.² Second, there is the neoclassical approach taken in the papers by James Tobin, David Levhari and Don Patinkin, Miguel Sidrauski, and Duncan Foley. They assume that: (a) the rate of capital formation is identically equal to planned savings, and (b) markets are always in equilibrium regardless of the rate of price change. Third, the Keynes-Wicksell approach, taken by Hugh Rose, Jerome Stein (1966, 1969), Keizo Nagatani, and Shō Chieh Tsiang assumes that: (a) prices are changing if, and only if, there is market disequilibrium and (b) there are independent savings and investment functions. During inflationary periods, when all demands cannot be satisfied, capital formation may differ from planned savings.

The aim of this essay is to discuss the following topics in monetary growth theory where attention is focused upon aggregative variables. What are the relations among the different approaches? What are the relative strengths and weaknesses of the various models? What are the substantive differences among models and which differences are inconsequential? To what extent is the distinction between outside and inside money crucial to the analysis? My main conclusion is that equally plausible models yield fundamentally different results. Recourse to models derived from the theory of rational be-

* I am indebted to Alvin Marty, Polly R. Allen, John Black, George H. Borts, Karl Brunner, Keizo Nagatani and Jürg Niehans for criticisms of earlier drafts. The author is professor of economics at Brown University. He thanks the National Science Foundation for financial assistance.

¹ This topic is discussed in Stein and Nagatani, but is ignored here.

² This is the subject of Mrs. Allen's Ph.D. thesis at Brown University.

havior does not resolve the ambiguities, since there are several different plausible utility maximizing models.³ This subject is now ready for sophisticated hypothesis testing. Since I am a proponent of the Keynes-Wicksell approach, it is possible that the other protagonists would not subscribe fully to the point of view expressed in this critical essay.⁴

Underlying the various approaches is a simple framework for the analysis of economic growth in a single sector economy. Assume that full employment always prevails and the labor force $N(t)$ grows exponentially at rate n . If technological change is of the Harrod-neutral type, then n may be interpreted as the growth of "effective" labor: the natural rate of growth plus the rate of Harrod-neutral technical change. Then output, $Y(t)$, depends upon the inputs of capital services, which are assumed to be positively related to the stock, $K(t)$, and of effective labor services, $N(t) = N(0)e^{nt}$. Output per unit of effective labor, $y(t) = Y(t)/N(t)$, is assumed to depend upon the ratio of capital per effective worker, $k(t) = K(t)/N(t)$, equation (1).⁵

$$(1) \quad y(t) = f[k(t)].$$

Investment per worker can be considered as the sum of two parts: the in-

vestment per worker required to maintain the current capital-labor ratio, nk , plus the time rate of change of the capital-labor ratio, Dk (where $D \equiv d/dt$). Only if investment per worker exceeds the amount required to provide the new workers with the existing ratio of capital per worker will the amount of capital per worker rise. From the definition of $k \equiv K/N$ we derive:

$$(2) \quad DK/N = nk + Dk.$$

Output per worker is divided between consumption per worker c and investment per worker DK/N . Therefore:

$$(3) \quad y = c + nk + Dk \quad \text{or}$$

$$(4) \quad Dk = (y - nk) - c.$$

Figure 1 graphs equation (4). The curve $y - nk$ represents the amount of output per worker available for consumption per worker plus the change in the ratio of capital to labor. The previous assumptions imply the shape of $y - nk$.

Assume that consumption per worker C_1 is positively related to the capital intensity, k . This assumption seems safe since a rise in k raises both output per worker and wealth per worker. At present, disregard the consumption function C_2 .

When the capital intensity $k(t)$ is below k_* then $y - nk$ exceeds c . Some output per worker is available to raise the capital intensity; and $k(t)$ rises. If $k(t)$ exceeded k_* , then the capital intensity would decline since c exceeds $y - nk$. Equilibrium capital intensity k_* is stable in Figure 1. At k^* , the output per worker available for consumption per worker and for a rise in the capital intensity is maximal. Obviously $y'(k^*) = n$ at this "Golden Rule" value.

In this exposition, it is assumed that there is no commodity money: i.e., all money is costless to produce.⁶ Moreover,

³ For lack of space, two major topics were omitted from the final version of this paper: (a) the optimal monetary policy in a growing economy and (b) the implications of different utility maximization models. They will be discussed in my forthcoming book. Relevant references are D. Cass and M. Yaari, P. Diamond, M. Friedman, M. Sidrauski, J. Stein, J. Tobin (1968), S. C. Tsiang, and H. Uzawa.

⁴ In many respects, the views expressed are consistent with Alvin Marty's penetrating analyses (1968, 1969) of the neoclassical model. There are also substantial differences between our points of view.

⁵ Not only is this production function assumed to be smooth and concave, but capital and labor are assumed to be essential for the production of output. Output per unit of capital Y/K would fall to zero, if the capital intensity grew infinitely large. Assume also that the marginal product of capital is larger than the growth rate when the capital-labor ratio is zero.

⁶ See Niehans for the case where there is a commodity money.

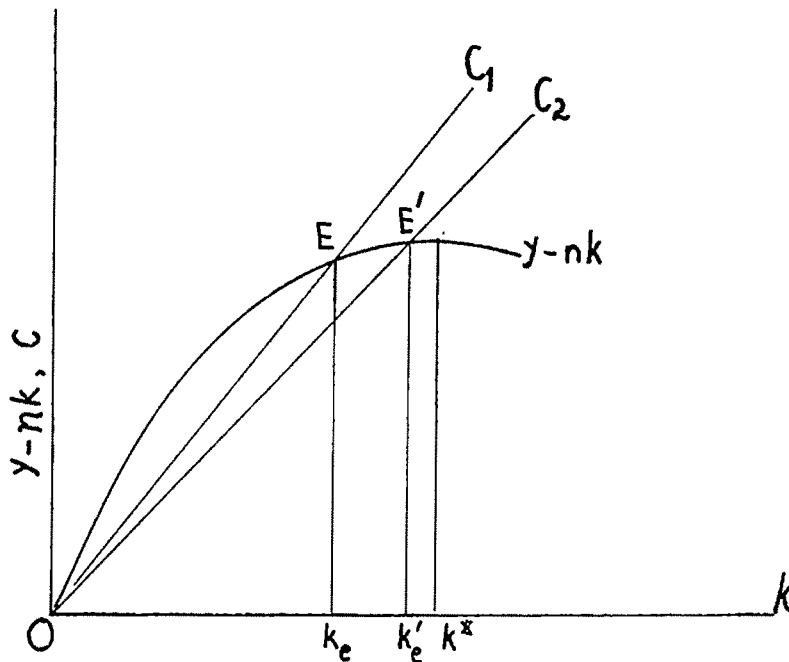


FIGURE 1. THE NEOCLASSICAL GROWTH MODEL. A RISE IN THE RATE OF MONETARY EXPANSION SHIFTS THE EQUILIBRIUM FROM E TO E' . $Dk = y(k) - nk - C[k + \theta L(k, \mu - n)]$.

for the sake of simplicity, only a single sector model will be considered.⁷

Monetary policy can affect the time profile of the capital intensity, as well as its steady state solution k_e , if it can shift either the consumption function or the net production function (defined as) $y - nk$. The neoclassical monetary growth model considers how these shifts can occur within a fully employed economy.

⁷ What is the role of money in such an economy? To avoid the complications of a multi-sector model, assume that the relative prices of a vector of goods are fixed. Thereby, we act as if the economy produced a single good. Insofar as the elements of the composite good are produced in different plants or firms, there is a need for a medium of exchange to avoid paying the workers in kind. The output produced by a worker consists of an element of the composite good, but his consumption consists of the composite good. Assume that each household, firm or plant is directly aware of the advantages of having the medium of exchange. The microeconomic foundation of the role of money (which is not interest bearing) in macroeconomic models is based upon the transactions demand and the precautionary demand for the medium of exchange. We shall assume that there is always less risk involved in holding money than there is in holding real capital.

Accordingly, a dichotomy is made concerning the role of real balances in the neoclassical model. First, under what conditions will variations in the rate of monetary expansion (and hence in real balances per worker) shift the consumption function? Second, under what conditions will the net production function $y - nk$ be affected by variations in the rate of monetary expansion? The first question is usually subsumed under the heading: real balances as a consumer good, and is discussed in part C of Section I. The second question is usually subsumed under the heading: real balances as a producer's good, and is discussed in part B below.

I. Neoclassical Monetary Growth Models

A. The Logical Structure

The neoclassical monetary growth model was first presented by Tobin (1965), and was a development made possible by the pioneering work of Robert Solow and T. W. Swan. It assumes that the rate of

capital formation is identically equal to planned savings and all markets are always in equilibrium, regardless of the rate of price change.

Monetary policy affects the consumption function in this model, even in the steady state.⁸ In this manner, monetary policy could affect the equilibrium capital intensity. Tobin assumed that consumption of goods per worker is a constant fraction of "disposable income" per worker. The latter is the sum of output per worker $y(k)$ and the increment of real balances per worker⁹ $D(M/p)/N$. The stock of money consists of the claims of the private sector upon the public sector which varies as a result of net transfer payments to or from the public. There is no other type of public debt or type of money. Assume that M grows exponentially and exogenously at rate μ ; and define the proportionate rate of change of the price level $\pi = D \ln p$, where p is the absolute price level. Then his consumption function is (5).

$$(5) \quad C/N = C[y(k) + D(M/p)/N]$$

$$(5a) \quad C/N = C[y(k) + (\mu - \pi)m]$$

where $m = M/pN$ is real balances per worker.

It would have been simpler, more general and more amenable to a dynamic analysis had Tobin used consumption function (6), which he suggested in an earlier paper in 1955. Let consumption (of

goods) per worker depend upon wealth per worker where wealth consists of real capital per worker, k , plus the real public debt per worker. It is not necessary to assume that the real public debt per worker is M/p . We could assume that it is $\theta M/p$, where θ represents the ratio of the nominal public debt to the stock of money M . Variable θ would be equal to unity if there were neither government bonds nor inside money. We may simply assume that θ is a positive constant. Then:

$$(6) \quad c = C/N = C[k + \theta m]$$

is the consumption function.¹⁰

Monetary policy will be able to shift the consumption function if it can vary m real balances per worker held, given the capital intensity.

The demand for real balances per worker is a function of transactions requirements per worker reflected by y ; the stock of the complementary asset per worker, k ; and the opportunity cost of holding real balances. Capital yields an expected return equal to its expected rent; and real balances yield an expected return equal to the negative of the expected rate of price change π^* . It is generally assumed that the expected rent is equal to the current level $r(k)$. The demand for real balances per worker can be written as equation (7).

$$(7) \quad m = L(k, \pi^*).$$

Clearly, $L_1 < 0$, since a rise in the expected rate of price change decreases the quantity of real balances demanded per worker at any given capital intensity. The sign of L_1 is positive since a rise in k raises the transactions demand for real balances and

⁸ Sidrauski derived a long-run consumption function which is independent of monetary influences. His conclusion was based upon two assumptions: (a) the economic unit is an immortal family maximizing utility over an infinite horizon and (b) the marginal rate of time preference is constant. The relaxation of either of these assumptions could change his results drastically.

⁹ Tobin assumed that total consumption C depended upon total real disposable income Y' . The latter is the sum of (a) total real output Y and (b) the change in the real value of the claims of the private sector upon the public sector $D(M/p)$, where M is outside money and there are no government bonds. Deflate by the size of the labor force N to obtain the measure of real disposable income per worker: $y(k) + D(M/p)/N$.

¹⁰ If the equilibrium $\mu - \pi = n$ is always positive, then it really does not matter in the steady state solution whether (5) or (6) is used. I shall use (6) rather than (5) in describing the neoclassical approach, since it is analytically more appealing and much simpler to use in a dynamic analysis. Tobin's 1965 paper was concerned with steady state solutions rather than a dynamic analysis.

also reduces the opportunity cost (the yield on real capital) of holding real balances.

The stability of the system is profoundly affected by the price expectations function. In the steady state, real balances per worker $m = M/pN$ are constant. Therefore, the price level eventually grows at a rate, π_s , equal to the growth of the money supply per worker. Let μ be the growth of the money supply. Then the steady state rate of price change π_s is equal to:

$$(8) \quad \pi_s = D \ln (M/N) = \mu - n.$$

A simple¹¹ price expectations function, which is both a stabilizing influence and is consistent with the steady state solution, is:

$$(9) \quad \pi^* = \mu - n.$$

Therefore, the consumption function is:

$$(10) \quad C/N = C[k + \theta L(k, \mu - n)];$$

and the time rate of change of the capital intensity is:

$$(11) \quad \begin{aligned} Dk &= y(k) - nk \\ &- C[k + \theta L(k, \mu - n)]. \end{aligned}$$

Figure 1 graphs this fundamental differential equation of neoclassical monetary growth theory.

A rise in the rate of monetary expansion shifts consumption function C downwards. Why? The rise in $\mu - n$ raises the expected rate of price change $\pi^* = \mu - n$. Thereby, there is a decline in the quantity of real balances demanded $L_2(k, \pi^*) < 0$ at any capital intensity $k > 0$. Since wealth per worker is $k + m$ (when $\theta = 1$ in most

models), the decline in $L(k, \pi^*) = m$ reduces real wealth per worker.

Consumption per worker is positively related to wealth per worker. Consequently, the rise in μ lowers the consumption function from C_1 to C_2 . The dynamics of this situation raises¹² the equilibrium capital intensity from k_s to k'_s . Monetary policy is able to affect the time profile of the capital intensity and its steady state solution, even if money is costless to produce.

B. Real Balances as a Producer's Good in a Neoclassical Economy

Two conclusions, which have been questioned recently, seem to emerge from the previous model. First, it would appear that inflation is conducive to economic development. A rise in the rate of growth of the money supply lowers the consumption function and raises the capital intensity. The average productivity of labor and the real wage are positively related to the capital intensity. Therefore, should developing countries, which want to raise real per capita income, be advised to inflate the growth of the money supply? Second, monetary policy was able to affect the time profile of the capital intensity because there was a real balance effect in the consumption function. If, however, all money were "inside" money, i.e., liabilities of a regulated but privately owned banking system, would there be a real balance effect in the consumption function? In that case, would monetary policy be able to affect the time profile of the capital intensity¹³ in this model?

Whether or not money is of the "inside"

¹¹ Alternatively, adaptive expectations function $D\pi^* = b(\pi - \pi^*)$ could be used. This function is not simple insofar as it adds a differential equation to the model. Moreover, the system will explode for a sufficiently high value of b , the adaptive expectations coefficient. If we assume that π^* is always equal to π , for all $D\pi^*$, then $D\pi^*/b = \pi - \pi^* = 0$ implies that b would be infinite. Such a model would not be stable.

¹² Stability requires that, at the equilibrium, the C function be steeper than the $y - nk$ function. At this stage of the argument, the C function could cut the $y - nk$ function above or below the Golden Rule level of k .

¹³ If the banking system uses a commodity money as reserves, then we must use something similar to Niehans' model.

or the "outside" type, real balances may be viewed as generating a productive service, complementary with labor and capital. If there were no medium of exchange, then the inefficiencies of a barter economy would result. Labor and capital would have to be diverted from the production of goods to their "distribution" in order to achieve the "double coincidence of wants." Firms would be established to act as brokers between potential buyers and sellers of goods and services; and the open book credit of those broker firms would undoubtedly develop as an inefficient money supply. What makes such a situation inefficient is that these firms use labor and capital to distribute goods and services which would otherwise be available for the production of goods and services. An explicit medium of exchange which is costless to produce increases the productivity of the economy by permitting a more efficient means of distribution and hence a greater rate of production of goods and services with given aggregate inputs of capital and labor.

There may be a real loss to society resulting from a reduction in real balances below a certain level. Either there must be more frequent payments, involving additional bookkeeping and other administrative expenses, or part of one's wage will be paid in kind entailing the use of some barter. For these reasons, aggregate output may be a monotonic nondecreasing function of real balances, regardless of whether money is of the inside or the outside type. We continue to assume that fiat money is used which is costless to produce.

If real balances are productive services, then the neoclassical model should be revised.¹⁴ Output per worker should de-

pend upon¹⁵ both capital per worker k and real balances per worker m , as described in equation (12).

$$(12) \quad y = y(k, m); \quad y_k > 0, \quad y_m \geq 0.$$

A limiting case that may occur with fully developed financial institutions is $y_m = 0$, i.e., an increment of real balances does not liberate any perceptible amount of resources. Assume however, that $y_m > 0$ for a sufficiently small m . Otherwise, why would money be used?

The private sector is assumed to allocate its wealth between the two assets: capital and real balances. In equilibrium, the net expected yields from each type of wealth will be equal. The anticipated return on capital is its expected rent, which is assumed to be equal to its current marginal product $y_k(k, m)$. The anticipated return on real balances has three components. First, there is the anticipated marginal product of real balances, which is assumed to be equal to its current level, $y_m(k, m)$. Second, there is the anticipated appreciation $-\pi^*$ in terms of its command over goods. Third, there is the "liquidity" yield of money $Z(k, m)$ which reflects the feeling that usually real balances are safer to hold real capital. Real balances are not treated here as a consumer good which yields utility directly. The liquidity yield reflects the price that asset holders are willing to pay, in terms of yield sacrificed, to hold an asset which may fluctuate less in real value. Assume that the liquidity yield is positively related to k and negatively

neglecting it, it is still in the tradition of 'solving' problems of monetary theory by metaphor instead of analysis. In fact, money is quite unlike the usual factors of production." We shall confine our attention to the aggregative level, and not try to answer Niehans' profound comment.

¹⁵ It is an oversimplification to assume that the production function $Y = Y[K, N]$, where N is effective labor, is unchanged during the process of economic development. For example, the allocation of resources between sectors has been improved during the process of economic development. See George Borts and Stein.

¹⁴ This aspect of the role of money has been stressed by Marty (1969), and Levhari and Patinkin. J. Niehans commented on this approach as follows: "While treating money as a 'productive service' may be better than

related to m , i.e., $Z_k > 0$ and $Z_m < 0$. It is, of course, possible that $Z(k, m)$, the relative liquidity of money compared to real capital, is negative; but that would be unusual. In any case, $Z(k, m)$ reflects the relative variances of the expected returns on capital and real balances.

Equilibrium requires that equation (13) be satisfied: the net yields of the two assets must be equal when the risk factor is taken into account.

$$(13) \quad y_k(k, m) = y_m(k, m) - \pi^* + Z(k, m).$$

Differentiate equation (13) and solve for dm in terms of dk and $d\pi^*$. Equation (14a) is derived; and equation (14b) is a more compact version of the same thing. If there are diminishing returns to substitution ($y_{ii} < 0$) and if the two inputs are complementary or independent ($y_{km} \geq 0$), then the denominator in equation (14a) is positive.

$$(14a) \quad dm = \frac{(y_{mk} - y_{kb} + Z_k)}{(y_{km} - y_{mm} - Z_m)} dk - \frac{1}{(y_{km} - y_{mm} - Z_m)} d\pi^*.$$

$$(14b) \quad dm = L_1 dk + L_2 d\pi^*.$$

Solving explicitly for the desired quantity of real balances per unit of effective labor, equation (14c) is derived. This is just the familiar portfolio balance equation.

$$(14c) \quad m = L(k, \pi^*); \quad L_1 > 0, \quad L_2 < 0.$$

Equation (14c) can be derived regardless of whether money is a liability of the public sector or of the private sector. The significant features of money, insofar as (14c) was concerned, are that it is a medium of exchange and a store of value.

An interesting and important question arises concerning the role of real balances as an argument in the consumption function. Continue to assume that consumption depends upon wealth. Should real balances be considered part of wealth if money is a liability of the public sector, but

excluded from a definition of wealth if money is a liability of the private sector? The question takes on added significance if real balances, regardless of who issues the money, affect the productive capacity of the economy.

In a barter economy, wealth (per effective worker) is k . What happens to wealth when an explicit medium of exchange is introduced and the inefficiencies of a barter economy are eliminated? Boris Pesek and Thomas Saving and Harry Johnson (1969) have asked whether there is an essential difference between outside and inside money. Does it make sense to say that wealth (per effective worker) is $k + m$ if the pieces of paper which serve as the media of exchange are liabilities of *all* of the people (the government), but that wealth (per effective worker) remains at k if the pieces of paper are liabilities of *some* of the people (a privately owned banking system)? Both types of money serve equally as well as media of exchange. The m in the production function $y(k, m)$ applies equally to outside and inside money. In view of this assumption, how should the consumption function be analyzed? We now discuss the implications of this issue for the neoclassical monetary growth model.

1. *Wealth consists of capital and real balances.* Suppose that consumers, in the aggregate, consider a dollar of outside money to be a larger increment to private wealth than a dollar of inside money, because there is a private liability associated with the latter but not with the former. Let M' represent outside money and M'' inside money. Nominal private wealth may be considered to be equal to $pK + \theta_1 M' + \theta_2 M''$, where $1 > \theta_1 > \theta_2 > 0$ represents the fraction of each type of money which the public considers to be private wealth. The conventional assumption is that $\theta_2 = 0$, but it is quite unnecessary for purposes of analysis. Let $M' = \xi M$ and

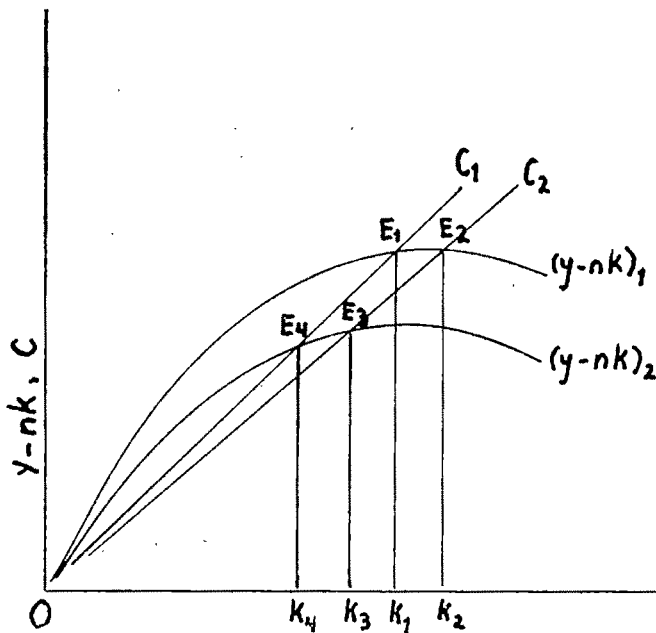


FIGURE 2. REAL BALANCES ARE PRODUCTIVE SERVICES. A RISE IN THE RATE OF INFLATION AFFECTS BOTH THE PRODUCTION FUNCTION AND THE CONSUMPTION FUNCTION.

$M'' = (1-\xi)M$ be the division of the total money supply between outside and inside money, respectively. Private nominal wealth is: $pK + [\theta_1\xi + \theta_2(1-\xi)]M$. If the weights θ_i and the division of the money supply ξ are constant, then $\theta = [\theta_1\xi + \theta_2(1-\xi)]$ will be constant. Real private wealth (per effective worker) would be $k + \theta m$. If all money were inside money, then real wealth per worker would be $k + \theta_2 m$; if all money were outside money, then real wealth per worker would be $k + \theta_1 m$.

Using consumption function (6), price expectations function (9) and production function (12), the basic differential equation for the time path of k can now be derived: equation (15). It is graphed in Figure 2.

$$(15) \quad Dk = y[k, L(k, \mu - n)] - nk \\ - C[k + \theta L(k, \mu - n)].$$

Equation (15) is more general than equation (11) for we allow for the possibility that $y_m > 0$, i.e., that the loss of real bal-

ances will adversely affect the production of output with given total supplies of labor and capital.

Let the rate of monetary expansion rise from μ_1 to μ_2 . When money is of the outside type, then a rise in μ entails a rise in the growth of the volume of net transfer payments. If money is of the inside type, then the liabilities of a privately owned banking system rise at a faster rate than before.¹⁸ There are now two effects. First, there may be an effect upon the demand for consumption per worker at any given capital intensity. A rise in μ raises the expected rate of price change π^* (equation (9)) and reduces the quantity of real balances demanded per worker ($L_2 < 0$ in equation (14)). If money is purely of the inside type, $\theta = \theta_2$, then wealth will be affected by $\theta_2 L_2$ and there will be an effect upon the consumption function equal to

¹⁸ See Section II A for a brief discussion of the determinants of μ in an inside money system, where a group received a franchise to print currency.

$C' \theta_2 L_2$. On the other hand, if money were of the outside type then there will be a larger effect upon wealth equal to $\theta_1 L_2$ and the consumption function will shift by $C' \theta_1 L_2$. The magnitude of the downward shift of the consumption function, given k , is:

$$\frac{\partial C}{\partial \mu} = C'[\theta_1 \xi + \theta_2(1 - \xi)]L_2 = C'\theta L_2,$$

and is described by the shift of the consumption function from C_1 to C_2 in Figure 2.

Second, the decline in the quantity of real balances demanded per worker may affect output per worker, i.e., $y_m L_2$ may be substantial (especially in an economy whose financial institutions are not fully developed). This means that the function $y - nk$ declines at the same time that the consumption function declines from C_1 to C_2 .

What will be the effect of a rise in the rate of monetary expansion upon the equilibrium capital intensity k ? It depends upon which effect dominates. In the original neoclassical model, the decline in the quantity of real balances per worker at any capital intensity did not affect output. This means that, in effect, a fully developed financial economy was considered where $y_m(k, m)$ was approximately zero. Hence, the $y - nk$ function did not shift. However, the economy contained outside money: i.e., real balances are considered part of wealth. The decline in m reduced wealth, and thereby lowered the consumption function. The rise in the rate of monetary expansion, in that case, raised the steady state capital intensity from k_1 to k_2 .

In the more general case analyzed here, it is not clear that inflation should be recommended as an aid to economic development. If real balances are highly productive (i.e., y_m is large), the decline in $y - nk$ could exceed the decline in C . The magnitude of the decline in C depends upon the division of the money stock between

inside and outside money (ξ) and the different weights attached to each form of wealth. If θ_2 is small, i.e., there is a weak real balance effect in the consumption function derived from inside money; and most of the money consists of inside money, i.e., ξ is close to zero: then there will be a small decline in the consumption function. Then a rise in the rate of inflation would lower the $y(k, m) - nk$ curve to $(y - nk)_2$ and lower the consumption function to C_2 . As a result of the rise in the rate of inflation, the equilibrium capital intensity would be lowered from k_1 to k_2 . In this formal, but general case, it was shown that inflation can lower output per capita and consumption per capita. A variety of results can be obtained from this model.

2. *Wealth consists of capitalized output.*

Instead of considering wealth as the stock of capital plus some function of real balances per worker, why not consider wealth as the capitalized flow of output? If a medium of exchange enables the economy to produce more output with the same expenditure of effort and the same rate of utilization of its equipment, then that medium of exchange is part of the nation's wealth. What difference does it make whether it is a liability of all of the people (the government) or of some of the people (a privately owned banking system)?

For simplicity, assume that (i) current output per worker $y(k, m)$ is expected to remain constant and (ii) the discount factor is the marginal product of capital $y_k(k, m)$. Then wealth per worker, a , is capitalized output per worker.

$$(16a) \quad a = y(k, m)/y_k(k, m);$$

and consumption per worker, c , depends upon wealth per worker, a .

With competitive pricing, the share of output earned by capital is β where:

$$(16b) \quad \beta = y_k(k, m) \cdot k / y(k, m) = \beta(k, m).$$

Therefore, wealth per worker is simply:

$$(16c) \quad a = k/\beta(k, m).$$

Empirically, the elasticity of substitution is equal to or less than unity which implies that $\beta_k \leq 0$. Therefore, the consumption function becomes:

$$(16d) \quad c = C[k/\beta(k, m)],$$

and c is positively related to k as before.

The remaining question is whether the share of capital β is affected by the quantity of real balances. Suppose that the share β were not affected by m . Under these assumptions, the growth model is described by equations (16), (14), and (8).

$$(16) \quad Dk = y(k, m) - nk - C[k/\beta(k)];$$

$$\beta' \leq 0.$$

$$(14) \quad m = L(k, \pi^*)$$

$$(8) \quad \pi^* = \mu - n$$

No distinction whatsoever is made between inside and outside money if they are indistinguishable as media of exchange and as productive services. There is no real balance effect in the consumption function if the share of output earned by capital is independent of the quantity of real balances. On the other hand, there is a real balance effect in the net production function.

No ambiguity exists in this model concerning the effect of a rise in the rate of monetary expansion upon the steady state capital intensity: it must *decline*.

A rise in the rate of monetary expansion lowers the quantity of real balances demanded at any given capital intensity: $L_2 < 0$. Real balances are productive services and therefore output per unit of effective labor changes by $y_m L_2 < 0$. The net production function declines from $(y - nk)_1$ to $(y - nk)_2$ in Figure 2. Consumption per effective worker is unaffected by a change in the quantity of real balances per worker, if capital's share of out-

put is independent of the quantity of real balances per worker. As a result of the rise in the rate of monetary expansion, the steady state capital intensity *declines* from k_1 to k_4 .

No ambiguity exists here. A rise in the rate of inflation *lowers* the steady state capital intensity, when wealth is measured as capitalized output. Inflation (in Figure 2) lowers the steady state output, and consumption, per unit of effective labor.

C. Real Balances as a Consumer Good

The analysis underlying equation (13) can be used to explain the derivation of equation (7): why real balances may be held even if they are not productive services on the margin (i.e., if $y_m = 0$). The answer would be that there is a liquidity yield $Z(k, m)$ derived from holding real balances rather than capital. This yield is positively related to transactions $y(k)$ and to wealth k ; and is negatively related to the quantity of real balances held. The optimum stock of real balances per worker will be held when the liquidity yield $Z(k, m)$ is equal to the opportunity cost $y_k(k) + \pi^*$. In this manner, equation (7) is derived when on the margin real balances are not a producer's good. But there does exist (m, k) such that $y_m > 0$.

Johnson (1967) and Levhari and Patinkin take a different approach and view the services of real balances as an addition to produced output, just as the services of owner occupied houses is part of national income. They view the services of real balances as a component of disposable real income which is consumed.

"Disposable real income" per worker in this framework is equal to output per worker $y(k)$ plus the real transfers per worker from the public sector $(\mu - \pi)\theta m$ plus the value of the services of real balances per worker. Most authors assume $\theta = 1$. How should the value of the services of real balances per worker be measured?

One approach, taken by Levhari and Patinkin, measures the value by the opportunity cost of holding real balances: the yield on capital less the yield on real balances. Assume that the expected yield on real capital is equal to its current rental r . What is the return derived from holding real balances? There may be a positive nominal interest rate paid on money i . Then, the expected yield on real balances is $i - \pi^*$, the nominal interest rate less the expected rate of price change. Therefore, the opportunity cost of holding real balances relative to capital is $r + \pi^* - i$, which is equal to the marginal utility of real balances as a consumer good. The value of the services of real balances is the quantity m multiplied by the opportunity cost $r + \pi^* - i$. Disposable real income, according to this approach, would be Y_d' (equation (17)).

$$(17) \quad \frac{Y_d'}{N} = y(k) + (\mu - \pi)m \\ + (r + \pi^* - i)m.$$

Another approach, taken by Johnson (1967), measures the value of the services of real balances, a component of disposable real income, as the integral under the demand curve for real balances per worker. Let $U(m)$ be the total utility per worker associated with the possession of real balances per worker of quantity m . Function U is monotonic nondecreasing. Then disposable real income per worker is (18).

$$(18) \quad \frac{Y_d''}{N} = y(k) + (\mu - \pi)m + U(m).$$

There are serious criticisms which can be levied at each approach. First, the usual definitions of real national income accounting are violated. Real output y is presumably measured in constant prices. Each item in the bundle of outputs is valued at constant prices, and the total is summed. This was not done for the ser-

vices of real balances. In equation (17), the real value of the services of real balances is measured in current prices since $r + \pi^* - i$ is not constant during the analysis. A paradox would arise if the demand for real balances had an inelastic section. A rise in the quantity m would be associated with a decline in $(r + \pi^* - i)m$, the total real value. The value of real output y is not measured this way. Why treat the value of the services of real balances in an asymmetrical way?

Equation (18) is not much better in this respect. The real value of the services of output consumed, a component of y , is not measured as an integral under a demand curve reflecting the total utility associated with the given quantity. Why measure the services of real balances as an integral under a demand curve? There is no justification for the asymmetry of treatment between currently produced output and the services of real balances, in a measure of disposable real income.

Second, there is a difficulty in interpreting the real balance effect in this approach. There are two components of consumption: the consumption of goods and the consumption of the services of real balances. These models generally assume that total real consumption of goods and real balances is a constant fraction $1 > c > 0$ of real disposable income. Therefore, the consumption of goods per worker C/N is derived from each definition of real disposable income.

Using the definition of disposable income in (17), the demand for consumption of goods per worker can be derived. Assume that $\theta = 1$ and there is no interest on money $i = 0$. Consider the steady state solution $\pi^* = \pi = \mu - n$. Then:

$$C/N = cY_d'/N - (r + \mu - n)m. \\ (19) \quad C/N = cy(k) + cnm - (1 - c) \\ \cdot (r + \mu - n)m.$$

Suppose population were constant, $n=0$. Then, the consumption function for goods is:

$$(19a) \quad C/N = cy(k) - (1-c)(r+\mu)m.$$

A *negative* relation exists between real balances held per worker m and the consumption of goods per worker. This is an unusual result.

Definition (18) yields the same paradox. This is described by equations (20) and (20a).

$$(20) \quad C/N = cY'_d/N - U(m), \quad U' > 0.$$

$$C/N = cy(k) + cnm - (1-c)U(m).$$

$$(20a) \quad C/N = cy(k) - (1-c)U(m),$$

$$\text{when } n = 0.$$

Again C/N and m are *negatively* related when population is constant.

When the services or real balances are regarded as a component of disposable income which yields utility directly, many paradoxes arise.

D. How Money Enters the Economy

All of the neoclassical models are very sensitive to the manner in which the money supply grows. It has hitherto been assumed in the outside money models that the money stock grows exclusively because of net transfer payments to or from the public. Therefore, wealth is affected by the rate of monetary expansion. Relax this assumption and suppose that the medium of exchange is a liability of the government which grows *exclusively* because it bears an interest rate i . Then, the growth of the money supply $D \ln M \equiv \mu$ is equal to i . Consider (for the sake of simplicity) the case where money is a producer's good, though similar results occur in the case where m yields utility directly. The steady state rate of price change $\pi = \mu - n$. Therefore, the expected real rate of interest on money is $i - \pi^*$

$= \mu - \pi^* = n$. The interest rate paid on money $i \equiv \mu$ offsets completely, in the steady state, the effects of changes in the rate of price change $\mu - n$.

When both k and m are held as producers' goods, their marginal returns must be equal. The yield on capital, $y_k(k, m)$, must be equal to the real yield on money $y_m(k, m) + n + Z(k, m)$. The demand for real balances per worker in the steady state must satisfy:

$$(21) \quad y_k(k, m) = y_m(k, m) + n + Z(k, m)$$

$$(22) \quad m = L(k, n), \quad L_n > 0$$

for the reasons discussed earlier. Real balances demanded per worker will be independent of monetary influences.

As long as real balances have some productivity $y_m > 0$, or are considered to offer a convenience or liquidity yield $Z > 0$, the marginal product of capital $y_k(k, m)$ must exceed the growth rate n . Otherwise, people would not wish to hold capital. If that occurred, $r < n$, capital decumulation would result; and the marginal product of capital would rise. There could occur no equilibrium with a positive capital stock if the marginal product of capital were less than the growth rate.

A simple institutional change, whereby the growth of the money supply results exclusively from the payment of interest on money, makes monetary policy neutral. Variations in the rate of monetary expansion cannot affect the steady state capital intensity.

E. The Instability Elements in the Neoclassical Model

In the neoclassical model, the supply of and demand for real balances are always equal: i.e., portfolio balance is always assumed to prevail. This assumption will lead to the instability of the model, unless frictions are introduced.

Suppose that the expected rate of price

change π^* were always equal¹⁷ to the rate currently experienced π . Then, the equality of the supply of and demand for real balances per worker implies that:

$$(23) \quad m = L(k, \pi); \quad L_2 < 0,$$

based upon equation (7) and the assumption that $\pi^* = \pi$. To induce people to hold a larger quantity of real balances, given k , the rate of price change must decline. Solving explicitly for π , we derive:

$$(24) \quad \pi = \pi(k, m); \quad \pi_2 < 0.$$

This is a condition for portfolio balance; but it does not explain what causes the price level to change.

The growth of real balances per worker $D \ln (M/pN)$ is the growth of the money supply per worker $\mu - n$ less the growth of the price level π . Since portfolio balance is always assumed to prevail, the growth of the price level is given by (24). We, therefore, derive:

$$(25) \quad D \ln M/pN \\ = Dm/m = \mu - \pi(k, m) - n.$$

Instability is quite apparent. Say that k is given and m is displaced above its equilibrium value. Will m return to its equilibrium? The answer is: No. Why? To induce people to hold the larger stock of real balances, the rate of price change must *decline*. Therefore, real balances per worker *rise* at a *faster* rate than before; and m deviates further away from equilibrium. Formally,

$$(26) \quad \frac{\partial}{\partial m} (Dm/m) \Big|_k = -\pi_2 > 0,$$

which is instability in the m direction. Equations (25), (4), and (6) imply saddle point stability.

This source of instability is the same as

¹⁷ Or, suppose it were positively related to the current rate.

the paradox that Hahn develops in his 1966 paper. The trouble with this neo-classical model is that: "... the price of money was changing because this was required for asset equilibrium and not because any reason was adduced why, in fact, it should change." (Hahn 1969, p. 183).

Stability of the neoclassical monetary growth model can be guaranteed¹⁸ if a sufficiently sluggish price expectations function is introduced. Price expectations function (9) was excellent for this purpose: it was constant at $\mu - n$, the steady state rate of price change. That is why we did not encounter any instability earlier. The adaptive expectation equation can also be stabilizing if the value of coefficient b in $D\pi^* = b(\pi - \pi^*)$ is sufficiently small.¹⁹ When portfolio balance is always assumed to prevail, frictions in the formation of price expectations are necessary for stability of the neoclassical model.

II. Keynes-Wicksell Monetary Growth Models

In sharp contrast to the neoclassical model is a set of models which assumes that: (a) prices are changing if, and only if, the goods market is not in equilibrium and (b) there are independent savings and investment functions.²⁰

My aim in 1966 and 1969 was to formulate a general macroeconomic model which contains money in an essential way regardless of whether it is inside or outside money.

¹⁸ Assume that in $y=f(k)$, $f'(0) = \infty$ and $f'(\infty) = 0$.

¹⁹ This assumption is made in Sidrauski, Foley and Sidrauski.

²⁰ This Keynes-Wicksell approach was taken independently in 1966 by Rose and Stein and is similar to early work by Hahn (1960, 1961). This method has been developed in 1969 in subsequent papers by Rose, Stein, Nagatani and Tsiang. As a result of his criticisms of the neoclassical model, Hahn (1969) also seems to lean in a Keynes-Wicksell direction. Cagan (1969) implicitly thinks along these lines in the short run, but not in the long run.

If the inputs of labor and capital were arbitrarily fixed, then it would look like a dynamic version of Patinkin's 1965 short-run aggregative model. Alternatively, this growth model would be the generalization of post-Keynesian macroeconomics to the problems of a growing economy: where the input of capital is endogenously determined and growing over time. Long-run equilibrium is nothing other than the steady state solution of the short-run dynamic model with endogenous capital. Or, the short-run dynamic model is a special case of the general growth model. It has been used both in a full employment context by Stein, Tsiang and Hahn and in the case of unemployment by Rose and Nagatani.

There are two essential differences between the neoclassical and Keynes-Wicksell monetary growth models. They concern the dynamics of price change and the existence of independent savings and investment functions.

A. *The Dynamics of Price Change*

The fundamental assumption, which distinguishes between the two classes of models, concerns the determinants of price change in continuous time. In the Keynes-Wicksell model prices are changing if, and only if, aggregate demand for goods differs from aggregate supply. Excess aggregate demand is planned consumption C plus planned investment I less output Y . Since planned saving is $Y - C$, excess aggregate demand is $I - S$. Assume that the rate of price change $\pi = D \ln p$ is proportional to excess demand (deflated, for convenience,²¹ by the stock of capital). The price change equation is:

$$(27) \quad Dp/p = \pi = \lambda(I/K - S/K).$$

Such an approach was taken by Tobin in

1955, but clearly not in his subsequent work on growth. He considered a model where money and capital were the only stores of value (i.e., there were no bonds). Portfolio balance was defined as a condition where the supply of, and demand for, real balances are equal. The demand for real balances per worker was similar to that used in Section I above. He wrote:

Portfolio balance is assumed to be the necessary and sufficient condition for price stability ($Dp=0$). If, instead, owners of wealth desire to hold more goods and less currency, they attempt to buy goods with currency. Prices are bid up ($Dp>0$). If they desire to shift in the other direction, they attempt to sell goods for currency ($Dp<0$). [p. 105]

The mathematical formulation of this statement would be equation (28). The rate of price change is positively related to the excess supply of real balances per worker. Use the same demand for real balances function as was used in the neoclassical model. Then:

$$(28) \quad Dp/p = \pi = h[m - L(k, \pi^*)].$$

According to the usual version of Walras' Law (in a two-asset model), the excess demand for goods $C + I - Y$ is equal to the flow excess supply of real balances. If the excess flow supply of real balances is positively related to the excess stock supply $m - L$, then equation (28) states that the rate of price change is positively related to excess aggregate demand per worker. Except for the arbitrary deflator, equations (27) and (28) are very similar.

An example of the implications of the dynamic Walrasian equation (27) or (28) will be helpful in seeing: a) the difference in point of view between the Keynes-Wicksell and the neoclassical monetary growth model; and b) why monetary nonneutrality can occur even if there is no real balance effect in the savings function.

Suppose that a group was given the franchise to print the medium of exchange

²¹ In Tsiang's adaptation of Stein's 1966 model, he deflates I and S by effective labor.

(currency).²² The output produced by the owners of the franchise (which shall be called the bank) is DM in nominal terms and DM/p in real terms.²³ Nominal output DM is exchanged for DM/p of goods with the nonbank public, which desires currency for the usual reasons.

The real value of the flow of output per worker produced by the bank is:

$$(29a) \quad DM/pN = (DM/M)M/pN = \mu m,$$

where m is real balances per worker and μ is the rate of monetary expansion. If people were always holding their desired stocks of real balances (an assumption which will be dropped very shortly):

$$(29b) \quad m = L(k, \pi^*),$$

where L is the quantity of real balances demanded per worker. Continue to assume that the expected rate of price change π^* is equal to the steady state rate of price change $\pi_s = \mu - n$. Then, the real value of the output per worker produced by the bank would be:

$$(29c) \quad DM/pN = \mu L(k, \mu - n).$$

If the marginal cost of producing currency is zero, then the rate of monetary expansion $\mu = \mu_0$ which will maximize the real revenue (per worker) per unit of time of the bank (DM/pN) will be such that:

$$(29d) \quad \mu_0 \frac{L_2(k, \mu_0 - n)}{L(k, \mu_0 - n)} = -1.$$

At the maximum profit rate of monetary expansion μ_0 , the demand for real balances per worker will have a unit elasticity. If there is a strictly positive marginal cost of currency creation then the rate of monetary expansion will differ from μ_0 . In either case, μ_0 is the rate of monetary ex-

pansion produced by the owners of the franchise. Its derivation is not essential to my argument.

The franchise owners produce the currency and try to exchange it for goods produced by the nonbank public. There are only two ways in which the bank can sell its output: by demanding goods in exchange for currency, or by demanding nonbank debt in exchange for currency. The first case will be considered here, and the second in the next section. When the owners of the franchise demand goods (in exchange for their currency) aggregate demand is directly affected. There is no reason why the nonbank public should always wish to give up the goods demanded by the owners of the franchise, at the given price level and capital intensity. This is the point of divergence between the Keynes-Wicksell and the neoclassical models. A market mechanism must be operative. What is its nature?

At any time the stock of money (currency) is $M(t) = M_0 \exp \mu_0 t$, based upon the desire of the bank to maximize its profits. The nonbank public desires to hold $p(t) L(k_t, \mu_0 - n) N(t)$ of nominal balances. Alternatively, the nonbank public wishes to hold $L(k_t, \mu_0 - n)$ of wealth per worker in the form of real balances; but $M(t)/p(t) N(t)$ of real wealth per worker in the form of currency is in existence. The gap $M(t)/p(t) N(t)$ less $L(k_t, \mu_0 - n)$ may arise because the franchise owner has been demanding goods at a faster rate than the nonbank public has been willing to give them up; or the latter may have decided to switch from real balances to goods. In either case, the gap $M(t)/p(t) N(t) - L$ represents an excess demand for goods.

If the rate of price change Dp is proportional to the level of excess demand, i.e., the difference between the stock of money in existence $M(t)$ and the quantity demanded by the nonbank public $p(t) L(\cdot) N(t)$, then equation (30) or (28) follows.

²² This example was inspired by Cagan in his unpublished 1969 paper, who must be absolved from any responsibility for the views expressed here.

²³ $D = d/dt$.

$$(30) \quad Dp = h \left[\frac{M(t)}{N(t)} - p(t)L(k, \mu_0 - n) \right],$$

where $M(t)/N(t) = (M_0/N_0) \exp(\mu_0 - n)t$.

To illustrate the importance of this key assumption, assume that k_t is at its equilibrium level $k(t) = k$. Then $p(t)$, the solution of differential equation (30), will eventually grow at rate $\mu_0 - n$. In the steady state:

$$(31) \quad \begin{aligned} \mu_0 - n = \pi = \pi^* \\ = h \left[\frac{M(t)}{p(t)N(t)} - L(k, \mu_0 - n) \right]. \end{aligned}$$

A permanent gap $(\mu_0 - n)/h$ will exist between the stock of real balances in existence $M(t)/p(t)N(t)$ and the quantity demanded $L(k, \mu_0 - n)$. It is this inflationary gap that is the driving force behind the rise in prices in the short run, in almost all dynamic models. Solution of differential equation (30) shows that a similar situation exists in the steady state. Long-run inflation, at a rate $\pi = \mu_0 - n = \pi^*$, implies long-run excess demand.

Expectations of the rate of price change enter via the demand for real balances; and we have assumed that people always act as if the expected rate of price change were equal to its steady state value. The long run (steady state) is consistent with the price expectations function. What must be stressed is that the steady state represents the asymptotic solution of the model; and is not a condition that is imposed upon the model regardless of its consistency with short run dynamics. The short run is a special case of the steady state.

The neoclassical model assumes that portfolio balance always exists, regardless of the rate of price change: i.e., $m = L(k, \pi^*)$. This crucial neoclassical assumption would make sense if (a) the speed of adjustment h were assumed to be infinite or (b) a discrete period, rather than a continuous time, analysis were used. In the latter case, no contracts would be binding

until all markets were cleared; and the market would then be closed until the following day. If equation (28) were used instead of equation (7), then very different results would be obtained from the neoclassical model. A real balance effect in the savings function would not be necessary for the rate of monetary expansion to affect the steady state capital intensity.

B. Independent Saving and Investment Functions

Suppose that the franchise owners try to sell their output in exchange for interest bearing debt; and they plan to use their future interest receipts to purchase consumer goods. The original sellers of the debt are business firms who would use the acquired currency to demand output in the form of investment goods. The investment demand function may be of the form described by equation (32). The desired proportionate rate of change of the ratio of capital per effective worker $I/K - n$ (where I is desired investment) is assumed to be proportional²⁴ to the difference between the expected yield on capital $r + \pi^*$ and the nominal rate of interest ρ on debt.

$$(32) \quad \begin{aligned} I/K - n &= r + \pi^* - \rho \\ &= r - (\rho - \pi^*). \end{aligned}$$

Alternatively we could say that the desired proportionate rate of change of the ratio of capital per unit of effective labor is proportional to the difference between the rent²⁵ per unit of capital and the expected real rate of interest $\rho - \pi^*$. If the rent per unit of capital r were equal to the real rate of interest $\rho - \pi^*$, then equation (32) states that firms would desire a constant ratio of capital per unit of effective labor. Hence, desired capital would grow at rate n . This equation is the generalization of the short-run investment functions to a growing

²⁴ For simplicity, the factor of proportionality is assumed equal to unity.

²⁵ Assume that the expected rent is equal to the current rent.

economy; and is consistent with marginal productivity theory.

The franchise owners use their newly produced money to purchase debt in the market, thereby lowering the nominal interest rate on debt. At this lower nominal rate of interest, there will be a rise in planned investment by firms since the expected yield on capital $r + \pi^*$ has risen relative to ρ the nominal rate of interest.

The decline in the nominal rate of interest ρ leads to an excess demand for goods: planned investment has increased without a corresponding decline in the demand for consumption! In the neoclassical model, on the other hand, there is no independent investment function; and planned investment is identically equal to planned savings by consumers. Monetary changes can only affect the real variables, in the neoclassical model, by working through the consumption function. It is not easy to reconcile the neoclassical model with the example of the franchise owners presented here.

There is an excess demand for goods resulting from the attempt of the franchise owners to sell their output. What will happen to the rate of capital formation? The Keynes-Wicksell model shows how the rate of monetary expansion, produced by the franchise owners or the monetary authority, will affect the capital intensity $k(t)$ in both the short run and in the long run (steady state). No real balance effect in the savings function is necessary for this result.

There are (at least) two versions of the Keynes-Wicksell model. In one, there is forced savings during inflationary periods: the actual rate of capital formation exceeds planned savings but is less than planned investment. In the other, the rate of utilization varies positively with the rate of price change.²⁶ Both versions imply mon-

etary nonneutrality, even if there is no real balance effect in the savings function. Only a sketch of the first version will be presented here.

C. A Heuristic Exposition of the Keynes-Wicksell Model

1. *The possibility of forced savings during inflationary periods.* During inflationary periods, the demand for output $C + I$ exceeds the capacity of a fully employed economy $Y = F(N, K)$ where N and K are the currently available input quantities, and the rate of utilization of K and N is constant. Since $C + I - Y = I - (Y - C) = I - S$ is positive, the question arises: how much of output will actually be allocated for consumption and how much will be allocated for investment. Clearly both consumers and firms cannot be satisfied simultaneously, during period of excess aggregate demand. Will the actual rate of capital formation be equal to planned savings, i.e., output less planned consumption? Or will the actual rate of capital formation be equal to planned investment? We assume that the actual rate of capital formation DK , during periods of excess aggregate demand, will be less than firms desire (I) but more than consumers plan to save (S). Neither investment plans nor consumption plans are fully realized in periods of excess aggregate demand. Everyone is partially frustrated. The actual rate of growth of capital DK will be such that $I > DK > S$. Specifically, assume that the actual rate of growth of capital DK will be a linear combination of planned savings and planned investment, equation (33).

$$(33) \quad DK/K = aI/K + (1 - a)S/K.$$

Coefficient a is institutionally determined such that $1 > a > 0$ during periods of excess aggregate demand. Even if there were perfect foresight that $C + I$ exceeded Y , not everyone could be satisfied. Which demands are frustrated and which demands

²⁶ See Council of Economic Advisers and *Studies by the Staff of the Cabinet Committee on Price Stability* for some evidence consistent with this assumption.

are satisfied has to be determined by the institutional structure. No such problem exists during deflationary periods when there is sufficient output such that consumption plans can be and are fully realized. Then $a=0$, and there is more capital formation than is desired by firms. Firms find that they have not been operating on their investment demand schedules since I differs from DK in periods of price change. During deflationary periods, the full employment assumption may be questionable. Hence, we shall confine our analysis to inflationary periods.

Using equation (27), which states that the rate of price change π is proportional to excess demand per unit of capital, the rate of capital formation is:

$$(34) \quad \begin{aligned} DK/K &= a\pi/\lambda + S/K; \\ 1 > a > 0 & \text{ when } \pi > 0 \\ a = 0 & \text{ when } \pi \leq 0. \end{aligned}$$

Forced savings per unit of capital $a\pi/\lambda$ occurs during inflationary periods and reflects the fact that consumers acquire less output than they planned. It is based upon the assumption that if $C+I>Y$, then consumers will find that their actual consumption-income ratio is less than their desired consumption-income ratio. To be sure, firms will find that the actual rate of capital formation is less than the planned rate.

The proportionate rate of change of the ratio of capital to labor Dk/k is DK/K less n the growth of effective labor. Variables π and S/K are endogenous²⁷ and contain k and m as arguments.

$$(35) \quad Dk/k = a\pi/\lambda + S/K - n.$$

The growth of real balances per worker Dm/m is equal to the growth of the money supply per worker $(\mu - n)$ less the growth of the price level π .

$$(36) \quad Dm/m = \mu - \pi - n.$$

²⁷ See Stein (1966, 1969) for this derivation.

In the steady state, k and m are constant at k_0 and m_0 respectively. Therefore: (a) capital and labor grow at exogenous rate n and (b) the equilibrium rate of price change π_0 is equal to the proportionate rate of change of the money supply per effective worker.

$$(37) \quad (DK/K)_0 = n.$$

$$(38) \quad \pi_0 = \mu - n.$$

Figure 3a describes the steady state when planned savings per unit of capital is primarily a function of output per unit of capital. Then, the S/K function is negatively sloped²⁸ and is not very sensitive to monetary disturbances. If there were price level stability, then there would be no forced savings. Planned savings per unit of capital would be equal to the growth of effective labor, at capital intensity k_0 .

What will be the effects of a rise in the rate of monetary expansion? Consider the inflationary case ($\mu - n > 0$) where forced savings occur. Figure 3a describes this situation. At capital intensity k_0 , capital and labor grow at the same rate; and the rate of capital formation is equal to planned savings. Inflation raises the rate of capital formation above the rate of desired savings, since it is assumed that the driving force behind inflation is excess aggregate demand, i.e., the difference between planned investment and planned savings. (Recall our example where the franchise owners are purchasing debt with their output of currency.) The curve DK/K shifts to $S/K + a\pi/\lambda$.

The capital intensity is raised above k_0 . As the capital intensity rises, the average productivity of capital Y/K declines. Therefore, planned savings per unit of

²⁸ For example, if $S=sY$ then $S/K=sY/K$. But Y/K is negatively related to the capital intensity k . Therefore, S/K is negatively related to k . In general, $S/K=S(Y/K)$, $S'>0$ and $Y/K=f(k)$, $f'<0$. Therefore, S/K and k are negatively related.

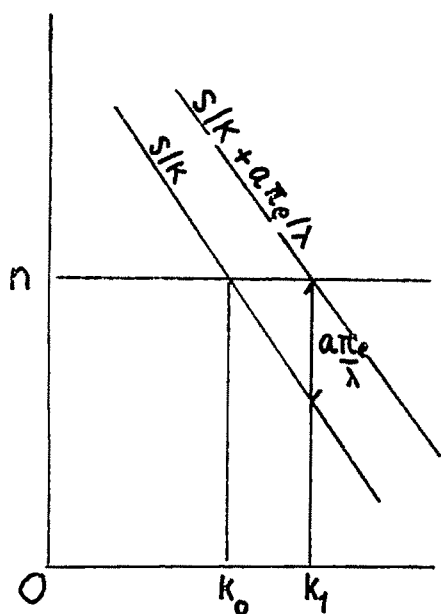


FIGURE 3a. A RISE IN THE RATE OF MONETARY EXPANSION RAISES THE STEADY STATE CAPITAL INTENSITY. INFLATION IS ASSOCIATED WITH EXCESS AGGREGATE DEMAND AND FORCED SAVINGS PER UNIT OF CAPITAL.

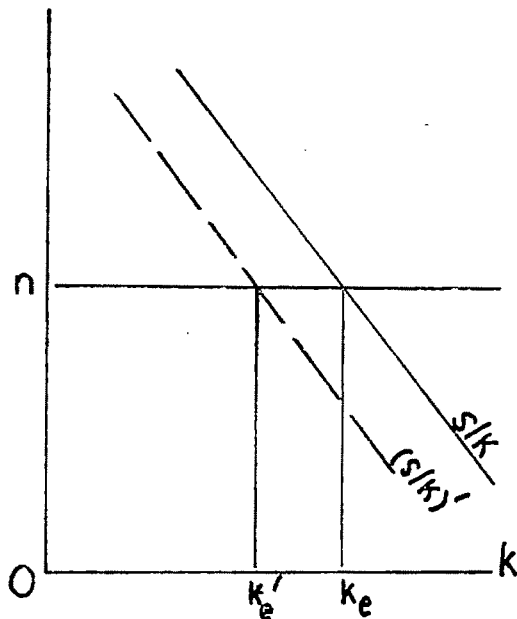


FIGURE 3b. A RISE IN THE RATE OF INFLATION LOWERS THE STEADY STATE CAPITAL INTENSITY. SAVINGS PLANS ARE ALWAYS REALIZED, BUT THE SUPPLY OF REAL BALANCES EXCEEDS THE DEMAND FOR REAL BALANCES.

capital also declines. Equilibrium is attained when planned savings per unit of capital declines by the amount of forced savings per unit of capital $a(\mu - n)/\lambda = a\pi_e/\lambda$. Here, the equilibrium capital intensity is positively related to the rate of monetary expansion. The original neo-classical result was obtained in a different manner. However, if S/K depended upon m and k , but $a=0$ (i.e., consumption plans are always realized) a rise in $\mu - n$ can lower the steady state capital intensity, in a dynamically stable model.

2. *Reverse results when savings plans are realized.* If saving plans are realized ($a=0$), then the growth of capital DK/K is equal to planned savings per unit of capital: equation (39).

$$(39) \quad n = S(k, m); \quad S_1 < 0, \quad S_2 < 0.$$

In the Keynes-Wicksell framework, the money market need not always be in equilibrium. If the bond market has a very

rapid speed of adjustment, then disequilibrium in the goods market is offset by disequilibrium in the money market. Equation (28), which was used by Tobin in 1955, states that the rate of price change is proportional to excess supply of real balances per worker. Rewriting (28), we obtain an expression for m the quantity of real balances per worker in existence, equation (40).

$$(40) \quad m = \frac{\pi}{h} + L(k, \pi^*).$$

In the steady state $\pi = \pi^* = \mu - n$. Using this relation in (40), and substituting (40) into (39) we obtain equation (41).

$$(41) \quad n = S \left[k, \frac{\mu - n}{h} + L(k, \mu - n) \right] \\ = \frac{DK}{K}.$$

When equation (41) is satisfied, then both

capital and labor grow at rate n . This is graphed in Figure 3b, where again S/K is negatively related to k , since $S_1 + S_2 L_1 < 0$.

A rise in the rate of monetary expansion $\mu = \pi_* + n$ is associated with a rise in the excess supply of real balances per unit of capital, equation (28) in the steady state. At any given capital intensity, the quantity demanded will decline, i.e., $L_2 < 0$. But the actual quantity of real balances per worker in existence will change by:

$$(1/h + L_2)\Delta\mu.$$

With a slow speed of response of price, the rise in the rate of monetary expansion will raise $M(t)$ faster than $p(t)$; and, therefore, m rises. If m rises, then S/K declines to $(S/K)'$ in Figure 3b. The decline in savings per unit of capital reduces the steady state capital intensity from k_* to k'_* .

Analytically, this result is obtained by differentiating (41) with respect to μ and solving for $dk_*/d\mu$.

$$(42) \quad \frac{dk_*}{d\mu} = (-S_2) \frac{(1/h + L_2)}{(S_1 + S_2 L_1)},$$

where $S_1 < 0$, $L_1 > 0$ and $L_2 < 0$. If $(1/h + L_2) > 0$, then a rise in μ raises m and lowers S/K . Hence, $dk_*/d\mu < 0$, i.e., equilibrium shifts from k_* to k'_* in Figure 3b. The neoclassical model assumes that h is infinite; but that is an arbitrary assumption and is counter to the usual practice in short-run dynamics. An important conclusion that emerges from the Keynes-Wicksell model is that a rise in the rate of monetary expansion can either raise, lower or keep the steady state capital intensity constant, in a dynamically stable system (see Stein 1969).

It is implicitly assumed that the rate of inflation is not sufficiently great as to interfere with the productivity of the economy; hence, the effects of inflation upon capacity output are ignored.

D. Conclusion and a Possible Synthesis of Neoclassical and Keynes-Wicksell Models.

Monetary growth theory can be made to yield a variety of qualitatively different results, depending upon which model is used. Both the neoclassical and the Keynes-Wicksell model imply that in the steady state: the expected and actual rates of price change are equal ($\pi_*^* = \pi_*$); and the actual rate of price change π_* is equal to the growth of the money supply per worker $\mu - n$. For expository purposes, assume that there is secular inflation in the steady state. The models differ with respect to the underlying *short-run* dynamic models; and this difference carries over to the steady state characteristics of the model. The steady state must be the asymptotic solution of the short-run dynamic model: it cannot be brought in as a *deus ex machina*.

Instantaneous equilibrium in all markets at all times is implied by the *neoclassical* model. As the excess demand curves shift, the equilibrium prices change immediately. During an inflationary process, everyone is always holding his desired real balances; and there is no excess demand for goods. Prices always rise instantly by the amount required to clear all markets. No sooner is the market cleared than the excess demand curve rises again; and a new and higher equilibrium price is established instantly. Stability in the neoclassical model requires that price expectations change sluggishly: e.g., a sufficiently low adaptive expectations coefficient must be assumed.

A very different approach is taken in the first version of the Keynes-Wicksell model. It is assumed that prices rise if, and only if, there exists excess demand. Expectations by themselves do not raise prices; but expectations working through excess demands produce price changes. Markets are not always cleared. At any time, rising prices imply disequilibrium, i.e., unsatisfied de-

mands. A Walrasian dynamic price adjustment process is postulated; and there is no presumption that prices move in precisely the manner postulated by the neoclassical model. Prices do not change too rapidly in response to changes in excess demand. A stable economy converges to a steady state in terms of variables deflated by the growing stock of capital or the growing size of the labor force. In this steady state, there is inflation which is anticipated $\pi_s^* = \pi_s = \mu - n$. Prices do not rise just because people expect them to do so, in a dynamically stable model. They rise steadily because there is permanent excess demand. A natural bridge is created between the short run and the steady state in this set of models. The Patinkin model is a special case of the Keynes-Wicksell model developed here.

A possible synthesis of the neoclassical and Keynes-Wicksell models could be made if the price change equation were:

$$(43) \quad \pi = Dp/p = \pi^* + \lambda(I/K - S/K).$$

In a perfectly competitive market, firms must take the market price as a datum; and production is adjusted accordingly. If there is excess demand (i.e., $I > S$) then there is some market bidding process that produces a rise in prices. Expectations of rising prices affect $I - S$ and only thereby affect the rate of price change. This was the rationale underlying equation (27). On the other hand, it is possible that firms may have some power to set prices, or that markets are organized by specialists who take the expected rate or price change into account in setting prices. The actual rate of price change may be the sum of two elements: the specialists' expectations of price changes and the actual state of excess demand, as described by equation (43).

If equation (43) were the correct price determination equation, then the Keynes-Wicksell model would describe the growing

economy outside of the steady state. As the economy approaches the steady state, π approaches π^* and I/K approaches S/K . Then (a) all markets would be in equilibrium and (b) the actual rate of growth of capital will be equal to S/K . Monetary policy would be able to affect the steady state capital intensity, with this synthesis, only if there were a real balance effect in the consumption (savings) function or in the net production function. The crucial question is: which is the correct monetary growth model?

REFERENCES

- G. H. Borts and J. L. Stein, *Economic Growth in a Free Market*, New York 1964.
- P. Cagan, "The Non-Neutrality of Money in the Long Run," *J. Money, Credit, Banking*, May 1969, 1, 207-27.
- , "The Revenue from Money Creation and its Disposition—A Theoretical Analysis," unpublished 1969.
- D. Cass and M. Yaari, "Individual Saving, Aggregate Capital Accumulation and Efficient Growth" in K. Shell, ed., *Essays on the Theory of Optimal Economic Growth*, Cambridge, Mass. 1967.
- P. Diamond, "National Debt in a Neoclassical Growth Model," *Amer. Econ. Rev.*, Dec. 1965, 55, 1126-50.
- D. Foley and M. Sidrauski, "Portfolio Choice, Investment and Growth," *Amer. Econ. Rev.*, Mar. 1969, 60, 44-63.
- M. Friedman, *The Optimum Quantity of Money*, Chicago 1969.
- F. Hahn, "The Stability of Growth Equilibrium," *Quart. J. Econ.*, May 1960, 74, 206-26.
- , "Money, Dynamic Stability and Growth," *Metroeconomica*, Aug. 1961, 13, 57-76.
- , "Equilibrium Dynamics with Heterogeneous Capital Goods," *Quart. J. Econ.*, Nov. 1966, 80, 633-46.
- , "On Money and Growth," *J. Money, Credit, Banking*, May 1969, 1, 172-87.
- H. G. Johnson, *Essays in Monetary Economics*, Cambridge, Mass. 1967.

- , "Inside Money, Outside Money, Income, Wealth, and Welfare in Monetary Theory," *J. Money, Credit, Banking*, Feb. 1969, 1, 30-45.
- D. Levhari and D. Patinkin, "The Role of Money in a Simple Growth Model," *Amer. Econ. Rev.*, Sept. 1968, 55, 713-53.
- A. Marty, "Some Notes on Money and Economic Growth," *J. Money, Credit, Banking*, May 1969, 1, 252-65.
- , "The Optimal Rate of Growth of Money," *J. Polit. Econ.*, July/Aug. 1968, 76, Part 2, 860-73.
- K. Nagatani, "A Monetary Growth Model with Variable Employment," *J. Money, Credit, Banking*, May 1969, 1, 188-206.
- J. Niehans, "Efficient Monetary and Fiscal Policies in Balanced Growth," *J. Money, Credit, Banking*, May 1969, 1, 228-51.
- D. Patinkin, *Money, Interest and Prices*, 2nd ed., New York 1965.
- B. Pesek and T. Saving, *Money, Wealth and Economic Theory*, New York 1967.
- H. Rose, "Unemployment in a Theory of Growth," *Int. Econ. Rev.*, Sept. 1966, 7, 50-58.
- , "Real and Monetary Factors in the Business Cycle," *J. Money, Credit, Banking*, May 1969, 1, 138-52.
- P. A. Samuelson, "An Exact Consumption-Loan Model of Interest with or without the Social Continuance of Money," *J. Polit. Econ.*, Dec. 1958, 66, 467-82.
- M. Sidrauski, "Rational Choice and Patterns of Growth in a Monetary Economy," *Amer. Econ. Rev.*, May 1967, 57, 534-44.
- R. Solow, "A Contribution to the Theory of Economic Growth," *Quart. J. Econ.*, Feb. 1956, 70, 65-94.
- J. L. Stein, "Money and Capacity Growth," *J. Polit. Econ.*, Oct. 1966, 74, 451-65.
- , "Rational Choice and Patterns of Growth in a Monetary Economy: Comment," *Amer. Econ. Rev.*, Sept. 1968, 58, 944-50.
- , "Neoclassical and Keynes-Wicksell Monetary Growth Models," *J. Money, Credit, Banking*, May 1969, 1, 153-71.
- , and K. Nagatani, "Stabilization Policies in a Growing Economy," *Rev. Econ. Stud.*, Apr. 1969, 36, 165-83.
- , "A Minimal Role of Government in Achieving Optimal Growth," *Economica*, May 1969, 36, New Series, 139-50.
- , "The Utility Yield of Real Balances in an Intergeneration Neoclassical Monetary Growth Model," unpublished 1969.
- T. W. Swan, "Economic Growth and Capital Accumulation," *Econ. Rec.*, Nov. 1956, 32, 334-61.
- J. Tobin, "A Dynamic Aggregative Model," *J. Polit. Econ.*, Apr. 1955, 63, 103-15.
- , "Money and Economic Growth," *Econometrica*, Oct. 1965, 33, 671-84.
- , "Notes on Optimal Monetary Growth," *J. Polit. Econ.*, July/Aug. 1968, 76, 833-59.
- S. C. Tsiang, "A Critical Note on the Optimum Supply of Money," *J. Money, Credit, Banking*, May 1969, 1, 266-80.
- H. Uzawa, "On a Neoclassical Model of Economic Growth," *Econ. Stud. Quart.*, Sept. 1966, 17, 1-14.
- U.S. Council of Economic Advisers, *Economic Report of the President*, with the annual report of the Council of Economic Advisers, Transmitted to the Congress, January 1969, Washington.
- , *Studies by the Staff of the Cabinet Committee on Price Stability*, Washington, January 1969.

Differential Tariffs, Negative Value-Added and the Theory of Effective Protection

By AUGUSTINE H. H. TAN*

The theory of effective protection (*EP*), as developed¹ by W. M. Corden (1966) and Harry Johnson, rests on the two critical assumptions of Leontief production coefficients and partial equilibrium.² Recent attempts by James Leith, Benton Massell, and Seija Naya and James Anderson to relax the assumption of fixed coefficients have all been in partial equilibrium terms.

In this article, we examine some of the implications for *EP* theory of allowing variable factor proportions and general equilibrium repercussions. It is shown that, under such conditions, (a) *EP* rankings do not necessarily provide any index of economic efficiency even when negative value-added occurs (Sections I and II); (b) *EP*

rates do not necessarily measure the maximum proportionate increase in factor rewards (Sections III and IV); and (c) *EP* rankings do not unambiguously indicate the direction of resource flows³ (Section V).

I. An Economic Interpretation of Negative Value-Added

The setting of our problem is an economy with a structure of differential protection such that consumption goods are heavily protected while material (intermediate) imports come in with little or no nominal protection (construed to be tariffs or quotas, or both). The economy has little or no monopoly power in international trade, so that external prices are fixed. Such an economy is characteristic of many less developed countries. (See McKinnon, p. 585.) We assume that the economy has adjusted to the tariff structure, and that the exchange rate is fixed. Let each industry, *i*, whose value-added we want to measure be in competitive equilibrium with a constant returns to scale production function, earning zero or normal profits. Each factor of production is paid its marginal product, assumed to be the same everywhere. Domestic product prices are fixed by the small country assumption and the given product tariff.

Let us then denote gross output of industry *i* by Z_i , and the corresponding production function as

* The material for this article is from a chapter of my doctoral dissertation. I am deeply indebted to Ronald I. McKinnon who, as teacher, scholar, and advisor exemplar, gave me his ideas so generously that he can rightfully claim joint authorship. T. N. Srinivasan gave me very helpful comments as did members of the Despres-McKinnon international economics seminar at Stanford University. I have also benefited from M. Aoki's comments.

The major support for the research came from the Agricultural Development Council, Inc., New York; further funds and facilities were provided by the Research Center in Economic Growth, Stanford University. I am in debt to all concerned. However, I retain sole responsibility for all deficiencies, errors, and heresies.

¹ V. K. Ramaswami and T. N. Srinivasan's paper is a further development of the point about resource allocation made in the present article. Corden's forthcoming paper and my January 1969 paper have recently put *EP* theory, with fixed coefficients, into a general equilibrium framework.

W. P. Travis' paper, although dealing with general equilibrium repercussions, does not make the points raised here.

² Although Corden (1966) raises some general equilibrium issues here and there, he does not have a consistent model for the purpose. The use of only one primary factor presents problems of knife-edged instability.

³ These findings throw into doubt other claims based on *EP* rates, such as comparative advantage rankings made, for example, by Bela Balassa and Daniel Schydrowsky.

$$(1) \quad Z_i = Z_i(B_1, B_2 \dots B_n; M_1, M_2 \dots M_r)$$

where the B_i 's refer to primary factors and the M_j 's are intermediate inputs. For simplicity, but without any loss of generality, we dispense with all but one intermediate input, M , and one domestic factor, B . Consider the conventional measure of value-added:

$$(2) \quad W_i = P_i Z_i - P_m M$$

Value-added, W_i , is defined as gross value of output, $P_i Z_i$, less the value of intermediate input, $P_m M$.

Since we are dealing with a small country, external prices denoted with asterisks are fixed; then, if no tariff is redundant:

$$(3) \quad P_i = (1 + t) P_i^*$$

and

$$P_m = P_m^*$$

The price of Z_i is one plus t times its world price, where t is the ad valorem tariff rate. Letting intermediate input in duty free, $P_m = P_m^*$, amounts to differential protection.

Consider then the procedure underlying empirical measures of value-added in world prices. Data on Z_i and M are obtained together with world prices, actual or imputed, and the following value-added in world prices, W_i^* , is computed:

$$(4) \quad W_i^* = P_i^* Z_i - P_m^* M$$

Recalling our price relationships (3) we can then write:

$$(5) \quad W_i^* = \frac{P_i Z_i - P_m M(1 + t)}{1 + t}$$

Comparing W_i and W_i^* from equations (2) and (5) we can readily see that W_i^* may be negative, depending on the magnitude of M and t , while W_i , domestic value-added may be positive. Therefore, value-added, measured in world prices

may be negative.⁴ The question then is: what is the economic inference that should be drawn from this empirical observation? Equation (5), if it is negative, appears to tell us that in terms of foreign exchange, we are throwing money down the drain if the industry is allowed to operate. However, the following analysis will demonstrate that this could be a highly misleading, in fact, erroneous interpretation. In the first place, equation (5) is *not* an equilibrium observation; it will be shown that efficient, indeed optimal, behavior of individual industries under constraints provided by a differential tariff policy may easily lead to measured negative value-added in world prices.

Consider the horizontal line AA in Figure 1. This line is defined by the material equivalent⁵ of a unit of output under world prices, given by the distance OA , where M assumes the value,

$$M^* = \frac{P_i^*}{P_m^*}$$

That is, M^* denotes the amount of material that may be obtained by exchanging one unit of output under world prices.

Similarly, the material equivalent of a unit of output under domestic prices is given by the distance OH (line HH), where M assumes the value;

⁴ The phenomenon of value-added turning negative when measured in world prices seems to be prevalent in many less developed countries using differential protection as a means of fostering industrialization. See Ronald Soligo and Joseph Stern. Various explanations of such observations range from the existence of gross economic inefficiency to the arguments by Giorgio Basevi (p. 150), and Paul Ellsworth (p. 401), that such results are absurd. The only paper offering the economic rationale, though without elaboration, is by Stephen Lewis and Steve Guisinger.

It should be noted that the purpose is not to deny the validity of arguments based on economic inefficiency; rather it is to highlight the adjustments necessitated by changes in the price regime and the consequences of variability in input coefficients.

⁵ I am indebted to Paul David's papers (1962, 1966) in developing the arguments around the concept of material equivalent of a unit of output.

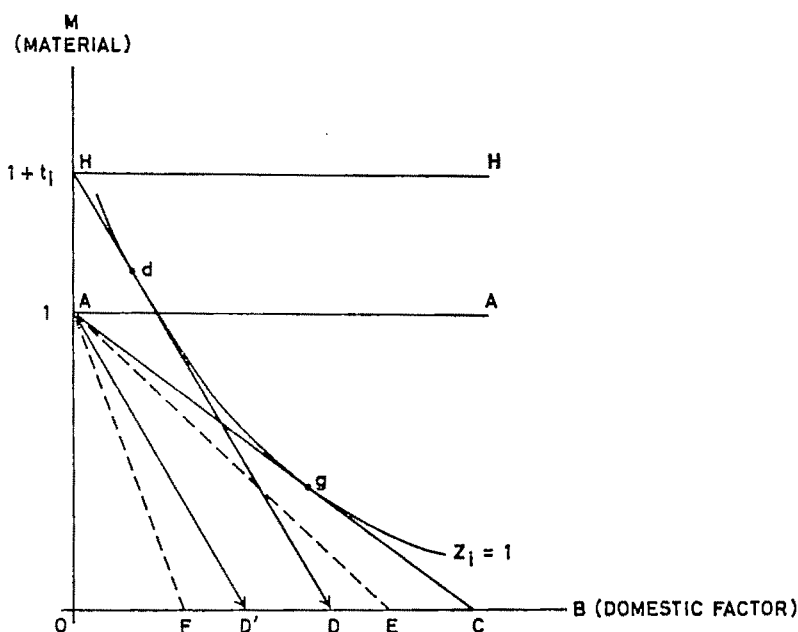


FIGURE 1

$$M^t = \frac{P_i}{P_m}$$

It should be noted that *neither* OA *nor* OH in Figure 1 denote *actual* material input. The introduction of the notion of material equivalents of a unit of output allows us to bring in output price levels into the analysis. Material input thus serves as a convenient numeraire.

Without loss of generality, we can normalize world prices to one:

$$P_m = P_m^* = P_i^* = 1$$

Then, assuming no redundancy in the tariff,

$$P_i = 1 + t$$

We can therefore write the material equivalents of a unit of output as:

- (6) $HH: M^t = 1 + t$, domestic price regime
 $AA: M^* = 1$, world price regime

Now, consider a point such as d in Fig-

ure 1. Point d gives us *positive* value-added in domestic prices because the amount of material used is less than HH . In terms of world prices, however, d represents negative value-added because the amount of material used is greater than the material equivalent of a unit of output in world prices, i.e. d is above AA . But what does the point d represent? In competitive equilibrium, *it is the resultant of the marginal equivalences between the unit isoquant and prices of output and of inputs, the latter defined with respect to the domestic price regime*. Therefore, point d is *not* an equilibrium observation with respect to world prices. The correct interpretation of the measured negative value-added as represented by point d depends on the level of analysis and type of policy being considered.

Case 1

Suppose the industry being considered is a small economic unit, such that a change

in policy, relevant only to this industry, has no factor price or other repercussions. It is in equilibrium at point d , with HD yielding the internal factor-price ratio. Under cost minimization, it is just breaking even with the given prices of output, imported material, and the domestic factor. Consider this industry in the absence of its output tariff protection. The resulting decrease in product price may be represented in Figure 1, by moving from HH to AA , the world price material equivalent of a unit of output. Since no factor-price changes are induced by the tariff removal, the relevant factor-price line becomes AD' , which is parallel to HD . This means, of course, that the industry must go out of business upon removal of its tariff protection, since it would then find it impossible to cover unit costs of production.

Case 2

Next, consider the case of a small industry such as in Case 1. However, the policy change contemplated is a revamping of the entire, or a substantial portion of the tariff structure. This will then mean that relative factor-material prices will be altered. (See Section III.) Suppose that along with the aggregate tariff changes, the tariff for our industry is removed. Then, as in Case 1, we move down from HH to AA . Factor-material price changes, however, could be in either direction, i.e., the domestic factor could become dearer or cheaper. In Figure 1, AF represents an increase in the price of domestic factor, B . Given the usual convexity properties of the isoquant, for the industry to be viable under the new tariff policy, the domestic factor has to be sufficiently cheaper. The price line AC , for instance, satisfies this requirement with equilibrium point g being attained. The line AE , though representing a cheapening of the domestic factor, does not allow unit costs to be covered. A fortiori, price line AF indicates nonviability without protection.

Therefore, in Case 2, with a small industry faced with a significant policy change, it is possible for the industry to be economically *viable without protection, even though it displays measured negative value-added in terms of world prices.*

Case 3

A third possible case is that of an industry which is significant enough such that a change in its rate of protection will generate some relative material factor-price change. A reduction of protection to this industry, with no changes in rates of protection in other industries, will lead to a contraction of the industry, decreasing its demand for the domestic factor. The industry will continue functioning so long as the domestic factor becomes sufficiently cheap to allow the industry to cover unit costs, e.g. point g in Figure 1.

Case 4

Next, consider the general version of Case 3 where the industry is big enough to affect domestic factor prices but such repercussions depend on changes in the rates of protection in other industries as well.

Again, in this case, as for Case 2, (with the exception that the factor intensity of the industry in Case 4 is relevant to factor-price changes), the domestic factor may become cheaper or dearer when the policy change of, say, removing the industry's tariff protection, is instituted. The relevant material factor price line may be AF , AE , or AC . AF and AE represent nonviability, as in Case 2. However, if the price line is AC , with equilibrium point g , the industry will continue functioning.⁶

What about those industries displaying positive value-added in world prices? Our analysis suggests that analogous considerations apply. Just as measured negative value-added in world prices should create

⁶ The analysis can be extended to cover questions of tariff redundancy, unimputed surplus, and a tariff on material imports without altering the conclusions.

no undue presumption of the industry being nonviable under world prices (except for Case 1), measured positive value-added in world prices should not create a presumption of viability under world prices, either (Case 1 *not* excepted). It is easily seen that for Case 1, measured positive value-added in world prices is no guarantee of viability under world prices—only when the point d coincides with the point g will the industry continue functioning under world prices. *But, this will happen only if the tariff was completely redundant initially.* Thus, for all practical purposes, and in the absence of technological change, all industries under Case 1 are nonviable when individual tariff protection is removed.

The analysis of measured positive value-added in world prices under Cases 2, 3 and 4 is analogous to the analysis of negative value-added, i.e. the industry may or may not be viable under world prices, no matter whether measured value-added in terms of world prices is positive or negative.

II. A Special Case: Leontief Technology

Up to now, we have assumed a variable-input technology. What about a Leontief technology, which excludes substitution between inputs? It can easily be shown that (i) all industries with measured negative value-added in world prices will be nonviable under world prices, for all four cases of Section I; (ii) all industries with measured positive value-added in world prices will be nonviable without individual tariff-protection (Case 1) except in the trivial case of complete tariff redundancy or unimputed monopoly gains; and (iii) in Cases 2, 3, and 4, all industries with measured positive value-added may, or may not, be viable under world prices, depending on relative material-factor prices, as in the case of variable input technology.

Thus the Leontief technology case appears as a special case of our earlier analysis. We may note that it is this special case which, implicitly for the most part,

underlies much of the discussion on negative value-added. The natural question arises: how appropriate or inappropriate is the assumption of a fixed coefficient technology? The proper response to this is, that it is an empirical question, and that to proceed to policy measures solely on an assumption may be highly misleading. Indeed, one can argue plausibly that the high incidence (see Soligo and Stern) of measured negative value-added in world prices creates a good presumption of a variable input technology *unless*, of course, all such industries did not exist prior to the introduction of the tariff structure.⁷ The argument is self-evident: if the industries existed, i.e. were viable, in the pre-tariff situation, then negative value-added in world prices cannot occur in the post-tariff situation if no factor substitution is allowed and if, of course, no technical retrogression occurs.

It seems much more satisfactory, therefore, to treat the Leontief technology case as no more than a special case and to rest the analysis⁸ on the safer ground of some variability in factor substitution.

III. Factor Price Adjustments

In the discussion of Cases 2, 3, and 4 in the preceding sections, one of the key elements was the behavior of the relative price of materials to domestic factors in going from one tariff policy to another. This question is ignored or else treated in partial equilibrium terms in the literature on value-added and on the theory of effective protection.

To highlight the nature of the problem we are dealing with, consider the following

⁷ Note that this argument applies only to Cases 2, 3, and 4 where factor price adjustments are admitted.

As a referee pointed out, this proposition will be difficult to test in less developed countries where many industries come into being after protection is granted.

⁸ I have extended the analysis to cover situations of nonconstant returns to scale, monopoly, factor market distortions, and learning effects, but space limitations prohibit their inclusion here.

simple model with two sectors; an importable dependent on labor and an imported input, and an exportable dependent on labor, imported input (material), as well as on a fixed factor, G . The introduction of the fixed factor is to prevent complete specialization since labor is the only other primary factor. We abstract from problems of tariff redundancy, monopoly, nonconstant returns to scale, and factor market distortions.

Let

$$(7) \quad \begin{aligned} Z_1 &= Z_1(L_1, M_1) \text{ importable} \\ Z_2 &= Z_2(L_2, G, M_2) \text{ exportable} \end{aligned}$$

Setting unit cost equal to price:

$$(8) \quad \begin{aligned} a_{1L}w + a_{1m}P_m &= P_1 \\ a_{2L}w + a_{2m}P_m + a_{2g}r &= P_2 \end{aligned}$$

where a_{ij} = amount of input j used per unit of output i .

Totally differentiating (8), dividing through by the respective output price, and letting

$$\hat{P}_1 \equiv \frac{dP_1}{P_1} \text{ and } C_{ij} \equiv \frac{a_{ij}\lambda}{P_i}, \quad \begin{aligned} i &= 1, 2; \\ j &= L, g, m, \end{aligned} \quad \text{as } \lambda = w, P_m, r;$$

$$\hat{a}_{ij} \equiv \frac{da_{ij}}{a_{ij}}$$

In general, the $\hat{}$ denotes proportionate change.

We get:

$$(9) \quad \begin{aligned} C_{1L}\hat{w} + C_{1L}\hat{a}_{1L} + C_{1m}\hat{P}_m + C_{1m}\hat{a}_{1m} &= \hat{P}_1 \\ C_{2L}\hat{w} + C_{2L}\hat{a}_{2L} + C_{2m}\hat{P}_m + C_{2m}\hat{a}_{2m} \\ C_{2g}\hat{r} + C_{2g}\hat{a}_{2g} &= \hat{P}_2 \end{aligned}$$

A necessary cost minimizing condition is:

$$\begin{aligned} wda_{1L} + P_m da_{1m} &= 0 \\ wda_{2L} + P_m da_{2m} + rda_{2g} &= 0 \end{aligned}$$

which become, in terms of C_{ij} , factor shares in unit costs:

$$(10) \quad \begin{aligned} C_{1L}\hat{a}_{1L} + C_{1m}\hat{a}_{1m} &= 0 \\ C_{2L}\hat{a}_{2L} + C_{2m}\hat{a}_{2m} + C_{2g}\hat{a}_{2g} &= 0 \end{aligned}$$

Therefore, (9) becomes:

$$(11) \quad \begin{aligned} C_{1L}\hat{w} + C_{1m}\hat{P}_m &= \hat{P}_1 \\ C_{2L}\hat{w} + C_{2m}\hat{P}_m + C_{2g}\hat{r} &= \hat{P}_2 \end{aligned}$$

Now consider a differential tariff structure such as we examined in our earlier discussion, with a tariff on the final importable and no tariff on the intermediate good or exports, i.e. $\hat{P}_1 > 0$, $\hat{P}_2 = \hat{P}_m = 0$.

From equation (11) we get immediately:

$$(12) \quad \hat{w} = \frac{\hat{P}_1}{C_{1L}} > 0$$

Since the denominator is the share of labor in unit cost of production of the final importable, it is necessarily less than one. Hence the wage rate rises *more* than proportionately when a tariff is imposed on the final importable. Conversely, going from a tariff situation to one of free trade, the wage rate will *fall* more than proportionately. Since P_m , the material price is fixed, this means that if the final importable industry does not operate on a Leontief technology, it may well display negative value-added in world prices but yet may be viable when allowed to adjust to world prices.

In terms of a more general model where both final goods industries depend on *two* domestic factors, say, capital and labor, as well as on imported material, wages and rents move in *opposite* direction, depending on capital-labor ratios. (A. H. H. Tan, 1969a.) Thus there is *both* substitution towards and away from material input. Nevertheless, in terms of our value-added measure, the *critical* point is that the two-way substitution is not exactly offsetting, i.e., we would observe (in a non-Leontief technology), a net substitution either away *from* or *towards* material input. The implication of this is that industries displaying

either positive or negative value-added may be viable under free trade if the net substitution is towards the domestic factors. If, instead, the net substitution is towards material inputs, then, in the absence of tariff redundancy or monopoly, the industries would go out of existence under free trade.

IV. *EP Rates and Maximum Factor Rewards*

The first claim of *EP* theory, that it provides a measure of the maximum proportionate increase in factor rewards permitted by a tariff structure, as compared to a free trade regime, can be readily disposed of. In the first place, the theory does not apply to cases where industries are viable in the tariff situation but not in the pre-tariff situation. However, even when industries are viable in both situations, there are difficulties.

Consider Case 1 of Section I: quite clearly, according protection *solely* to a small industry cannot affect the remuneration to domestic factors. What is likely to happen is either unimputed gains to the entrepreneur, who may or may not be considered a domestic factor, or the tariff is redundant.

What about Case 2 of a small industry, which is confronted with a broadly similar differential structure of protection across many such industries? The effect on domestic factor rewards must be analyzed from a general equilibrium vantage-point. The *EP* interpretation is clearly invalid for this case since, to a small industry, factor prices are given.

The *EP* interpretation fits our Case 3, the large industry case, *if no tariff changes elsewhere are being considered*. However, the latter is clearly not the premise of *EP* theory since more than one *EP* rate is considered. It is clear that *EP* theory deals with our Case 4, the large industry case with simultaneous changes in tariffs in many industries, *but in partial equilibrium*

terms. From a purely partial equilibrium standpoint, in order to attract additional resources, the industry must offer increasing factor prices. We note that *EP* theory implicitly imputes all the increase in product price (due to protection) to the return on the domestic factor(s).

V. *EP Rates and Resource Allocation*

The second goal of *EP* theory is to indicate the resource-allocative effects of a tariff structure. The measured rates are ordered on a scale through zero:

... if four activities producing traded goods can be ordered along a scale *A, B, C, D*, in ascending order of effective rates, we can say that output of *A* must fall and of *D* must rise and that resources will be pulled from *A* to *B* and from *A* and *B* to *C*; but without more precise information about production-substitution elasticities, we cannot say whether the outputs of *B* and *C* will rise or fall. [Corden 1966, p. 224]

The quotation above indicates that only the *ranking* of *EP* rates matter and that resources will move from sectors with lower *EP* rates to those with higher, since *EP* rates are interpreted to represent increases in the rates of remuneration to domestic factors. We shall demonstrate below that *EP* rates predict only a *particular* direction of resource flows whereas general equilibrium analysis, with variable input coefficients, indicates that flows contrary to *EP* prediction are possible. Two examples are provided: one, where there is a uniform tariff (subsidy) on final goods; and two, where one final sector has a higher nominal tariff than the other and there is no tariff on material imports. As will become apparent, the crux of the matter is that *EP* rates depend *not only* on nominal tariffs but *also* on value-added. The latter is a function of technology as well as relative output-input prices.

Consider again the simple model of Section III above. We shall present two

cases to show the ambiguity of *EP* theory. In *Case A*, the tariff on the final importable is equal to the subsidy on the exportable. In *Case B*, the tariff on the final importable is less than the export subsidy. There is no tariff on material imports in either case.

Case A

Under a uniform tariff and subsidy on the final goods,

$$\hat{p}_1 = \hat{p}_2 \quad \text{and} \quad \hat{p}_m = 0.$$

From equation (11) we get, again, that wages increase as in equation (12).

The rental on the fixed factor, (G), may, however increase or decrease:

From equations (11) and (12),

$$(13) \quad \hat{r} = \frac{\hat{P}_1(C_{1L} - C_{2L})}{C_{20}C_{1L}} \geq 0 \quad \text{as } C_{1L} \geq C_{2L}$$

To return to the mainstream of the discussion, we look next at factor endowment.

$$(14) \quad \begin{aligned} a_{1L}Z_1 + a_{2L}Z_2 &= L \\ a_{20}Z_2 &= G \end{aligned}$$

Totally differentiating equation (14) and converting to proportionate terms, letting

$$F_{ij} = \frac{a_{ij}Z_i}{\sum_i a_{ij}Z_i} = \text{fraction of supply of factor } i \text{ used in sector } j$$

and noting

$$\sum_i F_{ij} = 1,$$

we get:

$$(15) \quad F_{1L}\hat{Z}_1 + F_{1L}\hat{a}_{1L} + F_{2L}\hat{a}_{2L} + F_{2L}\hat{Z}_2 = 0$$

and

$$F_{20}\hat{Z}_2 + F_{20}\hat{a}_{20} = 0$$

or

$$\hat{Z}_2 = -\hat{a}_{20}$$

Next consider the production function Z_2 , from equation (7). By the linear homogeneity property, we can write:

$$Z_2 = Gf_2\left(\frac{L_2}{G}, \frac{M_2}{G}\right)$$

Hence

$$\frac{1}{a_{20}} = f_2\left(\frac{L_2}{G}, \frac{M_2}{G}\right)$$

Totally differentiating,

$$(16) \quad \hat{a}_{20} = -a_{20}\left[f_{2L}d\left(\frac{L_2}{G}\right) + f_{2M}d\left(\frac{M_2}{G}\right)\right]$$

We know the partial derivatives, being marginal physical products, are positive, i.e. $f_{2L}, f_{2M} > 0$. However, the total derivatives, $d(L_2/G)$ and $d(M_2/G)$, are dependent on relative input prices, P_m, r , and w . P_m is fixed; we know w always rises (equation (12)) but r may rise or fall (equation (13)). In addition, in a three-input world, two of the inputs can be complements. Hence, in general, the two total derivatives on the right-hand side of (16) may take either sign. Therefore \hat{a}_{20} may take either sign. Returning to equation (15) we therefore see that Z_2 may rise or fall, i.e. output of sector 2 may increase or diminish under a uniform tariff subsidy on final outputs.

Now let us look at what *EP* theory has to say. Suppose value-added in sector 2 exceeds that in sector 1. Then, since *EP* rates are given by:

$$(17) \quad T_i = \frac{t_i}{W_i^*}$$

where t_i = nominal tariff (subsidy) on sector i , it follows that W_i^* = equilibrium value-added in world prices in sector i .

This implies $T_1 > T_2$, *EP* rate of sector 1 exceeds *EP* rate of sector 2 even though both enjoy a uniform tariff/subsidy. According to Corden, therefore, resources will flow from sector 2 to sector 1. But we have

seen that this is not necessarily the case in general equilibrium with input substitution. *The Corden approach*⁹ totally ignores the general equilibrium role of material imports, factor-material substitution, and has no resource availability constraint.

Case B

For the second case, let $t_1 < t_2$, i.e. let the export subsidy exceed the final import tariff. Then $\hat{P}_2 > \hat{P}_1$. The wage-rate, w , rises as before (equation (12)). However, the rental, (r), on the fixed factor, (G) is now given by:

$$(18) \quad r = \frac{C_{1L}\hat{P}_2 - C_{2L}\hat{P}_1}{C_{1L}C_{2G}} \geq 0$$

as $C_{1L}\hat{P}_2 \geq C_{2L}\hat{P}_1$

Again, r can either rise or fall. The same arguments in Case A apply, i.e. sector 2 may either expand or contract.

Even though $t_1 < t_2$, *EP* rates may be such that $T_1 > T_2$ as in Case A, if value-added in sector 2 exceeds that in sector 1 sufficiently. Again, *EP* theory predicts a particular flow of resources, from sector 2 to sector 1, when the reverse flow is possible in general equilibrium and variable input proportions.

REFERENCES

- B. Balassa and D. M. Schydrowsky, "Effective Tariffs, Domestic Cost of Foreign Exchange, and the Equilibrium Exchange Rate," *J. Polit. Econ.*, June 1968, 76, 348-60.
- G. Basevi, "The United States Tariff Structure: Estimates of Effective Rates of Protection of United States Industries and Industrial Labor," *Rev. Econ. Statist.*, May 1966, 48, 147-60.
- W. M. Corden, "The Structure of a Tariff System and the Effective Protective Rates," *J. Polit. Econ.*, June 1966, 74, 221-37.
- , "Effective Protective Rates In The General Equilibrium Model: A Geometric Note," *Oxford Econ. Pap.*, July 1969, 21, 135-41.
- P. A. David, "Deflation of Value Added," *Rev. Econ. Statist.*, May 1962, 44, 148-55.
- , "Measuring Real Net Output: A Proposed Index," *Rev. Econ. Statist.*, Nov. 1966, 48, 419-25.
- P. T. Ellsworth, "Import Substitution in Pakistan: Some Comments," *Pakistan Develop. Rev.*, autumn 1966, 6, 395-407.
- H. G. Johnson, "The Theory of Tariff Structure with Special Reference to World Trade and Development," *Trade and Development*, Geneva: Institut Universitaire des Hautes Etudes Internationales, 1965.
- J. C. Leith, "Substitution And Supply Elasticities In Calculating The Effective Protective Rate," *Quart. J. Econ.*, Nov. 1968, 82, 588-601.
- S. R. Lewis, Jr. and S. E. Guisinger, "Measuring Protection in a Developing Economy: The Case of Pakistan," *J. Polit. Econ.*, Dec. 1968, 76, 1170-98.
- B. F. Massell, "The Resource-Allocative Effects of a Tariff and the Effective Protection of Individual Inputs," *Econ. Rec.*, Sept. 1968, 44, 369-76.
- R. I. McKinnon, "Intermediate Products and Differential Tariffs: A Generalization of Lerner's Symmetry Theorem," *Quart. J. Econ.*, Nov. 1966, 80, 584-615.
- S. Naya and J. Anderson, "Substitution And Two Concepts Of Effective Rate Of Protection," *Amer. Econ. Rev.*, Sept. 1969, 59, 607-11.
- V. K. Ramaswami and T. N. Srinivasan, "Tariff Structure and Resource Allocation in the Presence of Factor Substitution," Disc. Paper, No. 33, mimeo. Indian Statistical Institute, Planning Unit, New Delhi, July 1968.
- R. Soligo and J. J. Stern, "Tariff Protection, Import Substitution, and Investment Efficiency," *Pakistan Develop. Rev.*, summer 1965, 5, 249-70.
- A. H. H. Tan, "Differential Protection, Economic Indices and Optimal Trade Policies,"

⁹ The reference is to Corden's original contribution 1966.

unpublished doctoral dissertation, Stanford Univ. 1968.

———, (a) "Intermediate Products, Non-Traded Goods And The Samuelson-Stolper Theorem," Joint Memorandum No. 5, mimeo, Univ. Singapore, Oct. 1969.

——— (b), "Fixed Coefficients and The

Theory of Effective Protection in General Equilibrium," Joint Memorandum No. 4, mimeo. Univ. Singapore, Jan. 1969.

W. P. Travis, "The Effective Rate of Protection and the Question of Labor Protection in the United States," *J. Polit. Econ.*, June 1968, 76, 443-61.

A Geometric Treatment of Averch-Johnson's Behavior of the Firm Model

By E. E. ZAJAC*

The Averch-Johnson model of the regulated firm is now well-known, as is their conclusion that a profit-maximizing firm, regulated on fair rate of return, "operates inefficiently in the sense that (social) cost is not minimized at the output it selects," and, that "the firm adjusts to the [regulatory] constraint by substituting capital for the cooperating factor [labor]." This result has been named the "AJ effect," has been widely quoted, and has spawned a growing literature (see references to Bailey and Malone, Kahn, Klevorick, Trebing, and Westfield). It has also raised a fundamental issue: Regulation based on an objective of fair return may in fact be driving regulated firms to socially undesirable operations.

Unfortunately, Averch and Johnson carried out their analysis using the rather abstract tools of nonlinear programming and the Kuhn-Tucker theorem. Its detailed understanding has thereby been denied a large number of persons concerned with regulation to whom those tools are foreign and uncomfortable. This had led, in particular, to some misunderstanding about the relationship of their result to the general overcapitalization issue. "Overcapitalization" is used in regulation to

refer to several phenomena. Some examples are: inclusion of items in the rate base whose classification as capital investment is debatable ("inflating" or "padding" the rate base); addition of useless or idle plant (also called rate base inflating or padding); addition of plant at a rate faster than called for by an optimum investment strategy. The excessive substitution of productive capital for labor discussed by Averch and Johnson is a still different form of overcapitalization. Recent writings on regulation often confuse the various forms and give the impression that the Averch-Johnson paper has mathematically demonstrated a tendency of regulated firms toward all forms of overcapitalization (see papers edited by H. M. Trebing).

To clarify issues, a simple geometric analysis of the AJ model is given which uses no advanced mathematics. It is hoped that it will help show precisely what the model does and does not imply. The geometric analysis is followed by a discussion of the AJ result as descriptive economics. First, the behavior of the firm and its regulators is examined. With regulatory lag taken into account, the profit maximizing firm's best strategy is not clear; it may be to operate with minimum-cost inputs rather than the overly-capital intensive inputs implied by the AJ effect. Second, profit maximization is replaced by maximization of stockholders' rate of return (return to equity capital), another and perhaps more tenable model of management's objective. Here it is found that, without further assumptions on manage-

* The author is head of the economic analysis and graphics research department of the Bell Telephone Laboratories, Inc. He is indebted to W. J. Baumol, O. Eckstein, D. A. Kendrick, and F. W. Sinden for helpful comments on earlier drafts of this paper. The views presented here are solely those of the author and in no way express the opinions of the Bell Telephone Laboratories or of the American Telephone and Telegraph Company.

ment behavior, the firm is not driven to a unique capital-labor mix. Two further plausible patterns of management behavior are assumed. One again leads to minimum-cost inputs, the other to overly-labor intensive inputs. Third, an attempt is made to model the effect of the labor constraints to which firms are subject. The result is a model which is symmetric to the original AJ model, with the profit-maximizing firm tending toward overintensive labor use. Finally, rate base expansion by acquisition of nonproductive capital is examined geometrically to show how it differs from the AJ effect.

I. The Averch-Johnson Model and Its Geometric Solution

Consider a firm producing a single good, q , with two factor inputs, K (capital) and L (labor). Its output q is given by a production function, $q=q(K, L)$ with the price of q given by the inverse demand function, $p=p(q)$. The firm's per unit costs of capital and labor are given by i and w so that its total costs are $iK+wL$ and its profit, π , is

$$(1) \quad \pi = pq - (iK + wL).$$

Since p and q depend on K and L , the profit, π , also depends only on K and L . In a K, L, π coordinate system, the function $\pi=\pi(K, L)$ can be visualized as a surface, the "profit hill," spanning the K, L plane (Figure 1). It is assumed that for each K, L there is a single value of π , that the profit hill has a single peak, and that the surface continually falls away from the peak, so that the peak is the only point which is tangent to a horizontal plane.¹

¹ Similar assumptions are implicit in an analytic derivation of the AJ result. For example, multiple peaks lead to the possibility of multiple values of the Lagrange multiplier. These correspond to multiple, local constrained profit maxima, or minima. Such mathematical complications obscure the economic issues, and, as in the AJ paper, are avoided here.

Regulation may limit the firm so that it cannot operate at the very top of the profit hill. To model regulation, Averch and Johnson neglect depreciation and take K , the amount of capital the firm uses, to be its rate base. The money earned to be applied to the rental of capital is revenue minus labor expense: $pq-wL$. This divided by the rate base gives the rate of return to capital. Regulation requires that this rate be no greater than a fair rate of return, f . Thus, symbolically, the AJ model of the regulatory constraint is

$$(2) \quad \frac{pq - wL}{K} \leq f.$$

It is more convenient to rearrange (2) by multiplying by K and to use equation (1) to get

$$(3) \quad \pi \leq (f - i)K.$$

Averch and Johnson assume that the allowed rate of return f is greater than the average cost of capital i . The regulated firm is then allowed a positive (excess) profit, but by the constraint (3) this must be no greater than the product of the excess rate of return, $(f-i)$, and the rate base K .

The constraint $\pi \leq (f-i)K$ requires that the firm operate below or on the plane $\pi = (f-i)K$ in the K, L, π coordinate system. This plane can be visualized as a door hinged on the L -axis and swung upward from the K, L plane. The greater the value of the fair rate of return f , the higher the constraint plane is swung above the K, L plane (Figure 1).

In the case of interest, the constraint plane slices a prohibited bump off the top of the profit hill. The facts of demand and production put the firm on the profit hill but, except for uninterestingly high f 's, the constraint plane keeps the firm from attaining the hill's top. At the same time, the point of highest profit on or below the constraint plane is obvious. Because the

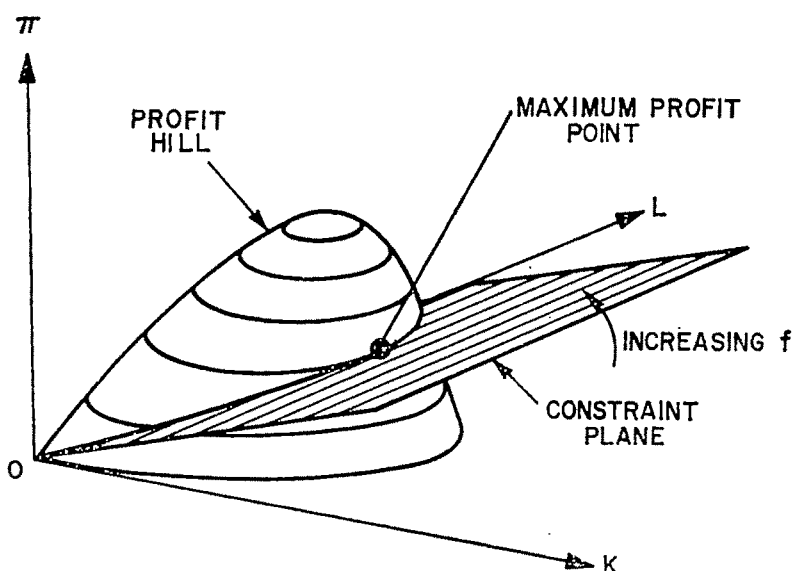


FIGURE 1
Profit Hill and Constraint Plane

constraint plane is hinged on the L -axis, *maximum profit occurs at the point of maximum K along the intersection of the constraint plane and the profit hill* (Figure 1).

To complete the analysis of the AJ model, it remains to be shown that at the K_{\max} point, the firm will operate inefficiently with overintensive capital utilization. For a fixed output (along an isoquant), efficient operation occurs at that allocation of capital and labor resources which results in minimum cost of production. Likewise, since revenue is fixed along an isoquant, the efficient point represents maximum profit to the firm along that isoquant. The π_{\max} point at the top of the profit hill (Figure 1) is the maximum of all fixed-output profit maxima, and hence lies on the locus of efficient points (expansion path). Furthermore, in the simple case considered here of a profit hill which has a single peak and which is nowhere tangent to a horizontal plane, the fixed-output profit maxima continually decrease as one moves away from π_{\max} in either direction

along the locus of efficient points.

Consider then the projection onto the K, L plane of the intersection of the constraint plane and the profit hill (Figure 2). This will be called the *constraint curve*. It encloses the region beneath the prohibited bump (shown shaded in Figure 2)². If the firm operates within the shaded region, it violates the regulatory constraint, while operation at the K_{\max} point results in maximum profit under regulation. The isoquant through K_{\max} cannot intersect the locus of efficient points *outside* of the constraint curve, as at point P in Figure 2 (a). For then, since π_{\max} is *inside* the constraint curve, the point P' at the intersection of the locus of efficient points and the constraint curve would lie between π_{\max} and P (Figure 2(a)). Inasmuch as profit continuously decreases along the efficient point locus away from π_{\max} , profit at P'

² The constraint curve need not be tear-drop shaped as shown. It may, for example, be heart shaped, with the point of the heart at the origin, or kidney shaped. The particular shape does not affect the conclusions.

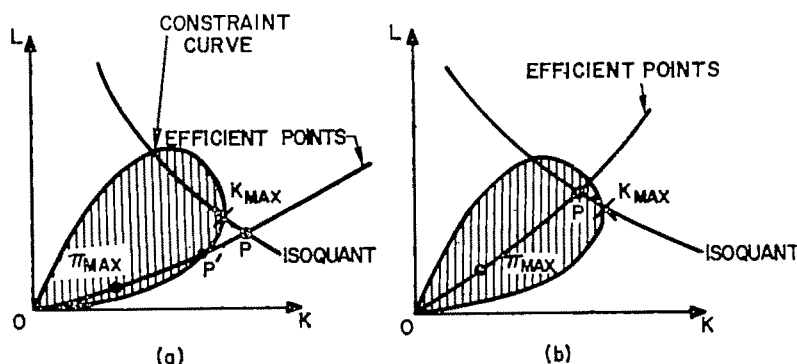


FIGURE 2

- (a) Impossible Location, Outside the Constraint Curve, of the Efficient Point P for the Isoquant Through K_{max}
 (b) Only Possible Location of P , Inside the Constraint Curve

would be greater than at P , and hence still greater than at K_{max} . But this would contradict the fact that K_{max} is the maximum profit point along the constraint curve. Likewise, coincidence of the K_{max} and efficient points can also be ruled out. Profit is a maximum at K_{max} as one traverses the constraint curve and at the efficient point as one traverses an isoquant. Hence, if the two coincide, profit is a maximum along two directions through the point of coincidence. This implies that the coincident point is tangent to a horizontal plane, which by the original assumption occurs only at the profit hill's peak. In other terms, K_{max} will also be an efficient point only when the constraint plane slices through the top of the profit hill.

Hence, the efficient point for the output curve passing through K_{max} must be inside the constraint curve as in Figure 2(b). To maximize profit under regulation, the firm should operate at point K_{max} in Figure 2(b). But the same amount of output could be obtained at lower cost if the firm were to move to efficient point P in Figure 2(b). Since a higher-than-necessary cost to the firm means the inefficient use of resources, society is the loser (so is the firm because it could increase its profits by moving to

P). Fair rate of return regulation thus would appear to drive the profit-maximizing firm to an operating point which is undesirable to society. This, then, is the oft-quoted AJ result, obtained originally by the application of the Kuhn-Tucker theorem rather than by the geometric arguments given here.

II. Assumptions About the Behavior of the Firm and Its Regulators

A key assumption in the Averch-Johnson model is that regulation results in an allowed rate of return, f , which exceeds the cost of capital, i . If the regulators succeed in setting $f=i$, then the firm's profits are at best zero (the constraint plane is horizontal); there is no question of maximizing excess profit under constraint and no AJ effect. But aside from this key assumption, the Averch-Johnson analysis also rests on the usual assumptions of the classic theory of the firm—perfect knowledge of markets and production, profit maximization, and operation at a static equilibrium. In addition, the analysis implicitly assumes a behavior of both the firm and its regulators which will allow the firm to operate at the K_{max} point. Since, to some extent, these assumptions are violated in the real

world, it is instructive to consider situations where some of them do not hold.

Perhaps most critical are the assumptions of a static equilibrium and the assumption that the profit maximizing firm will have enough information to operate at the K_{\max} point. For example, good engineering may allow the firm to operate with close to minimum cost technology. On the other hand, demand information is generally difficult and expensive to obtain and uncertain, thereby making difficult the determination of the shape of the profit hill. To illustrate the difficulties of deciding how the firm will act in the face of these as well as dynamic factors, let us consider two possible strategies for the profit maximizing firm:

Strategy I

The firm constantly strives to operate at lowest costs.

In following Strategy I, the firm always operates somewhere on the locus of efficient points, and the capital overintensiveness of the AJ effect does not occur. If it were at an efficient point, say P_1 in Figure 3, which violated the regulatory constraint, it would presumably incur a price cut. Successive price cuts would finally put the firm at P_{final} in Figure 3, the efficient point of minimum price and maximum output in the area enclosed by the constraint curve.

Vertical lines, $K = \text{constant}$, are lines of constant π in the constraint plane, $\pi = (f-i)K$. Hence, profits everywhere along the heavy arc of the constraint curve in Figure 3 are higher than at P_{final} . For the long term, a strategy that gets the firm, if not to K_{\max} , at least to this arc, might be preferable. But to get to K_{\max} , the firm must either have knowledge of demand in order to determine the profit hill and the K_{\max} point, or it must follow a procedure, as in Strategy I, that searches out K_{\max} .

Since the former has been assumed difficult and uncertain, let us consider the latter possibility in a Strategy II.

Strategy II

After a price and corresponding output level have been established, the firm adjusts its K, L mix so as to earn just the allowed rate of return. That is, if the firm is at P_1 in Figure 3, it moves along the constant output curve to C_1 , at the intersection with the constraint curve. Now it asks for a small price decrease and adjusts K, L to put itself at the intersection C_2 in Figure 3. If the profit at C_2 is higher than at C_1 , it asks for another small price decrease, and so on until a price decrease is followed by a profit decrease (which may occur after the first step). By this process, the firm is assured of operating somewhere along the favorable arc, where profits exceed those at P_{final} , and if it moves with small steps, it will operate very near to K_{\max} .

Thus, Strategy II searches out the K_{\max} point. But which strategy, I or II, yields the higher profit? Clearly this depends on timing factors that have not been considered. In Strategy I, the total profits in the move from P_1 to P_{final} depend on the duration of lags between attainment of excess rates of return and corrective action (regulatory lag). Likewise, in Strategy II, profits summed over time depend on the size and timing of the steps from C_1 to K_{\max} . Furthermore, it is not clear that Strategy II is preferable for the long term. The time necessary to get clear market response and the regulatory approval required to settle on the size of the corrective step might be of the order of months or even years in both Strategies I and II—time during which technology and market changes can cause radical shifts in both P_{final} and K_{\max} . Hence, the dynamic factors and imperfect information appear

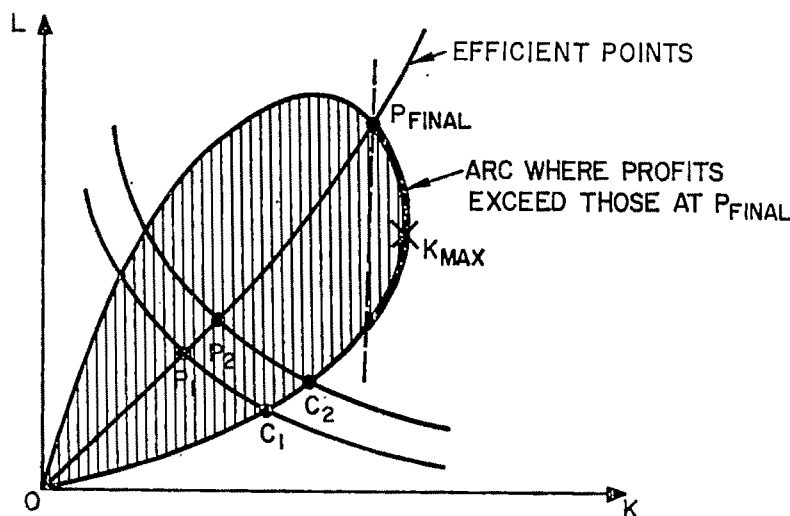


FIGURE 3
Possible Operating Points for the Firm

to make it difficult to say a priori what the firm's strategy for optimum profits ought to be. Likewise, without empirical studies, they would appear to make it difficult to assess the point of the K, L diagram at which the firm is to be found and the magnitude of the AJ effect.

III. Maximization of Rate-of-Return to Equity Instead of Profit

Averch and Johnson assume that the firm wishes to maximize total profit. Suppose instead that the goal of the firm's managers is to reward the owners (stockholders), that is, to maximize net return per dollar of stockholder investment.³

Assume that a fraction f_d of the firm's capital K is debt and a fraction f_e is stockholder's equity ($f_d + f_e = 1$), and that these have constant values.⁴ Then the rate of

return to stockholder equity, r_e , is

$$(4) \quad r_e = \frac{pq - wL - i_d f_d K}{f_e K}$$

where i_d is the bond rate, and $f_d K$ and $f_e K$ represent the amounts of debt and equity capital respectively. As before, the regulatory constraint is

$$(5) \quad \frac{pq - wL}{K} \leq f.$$

The managers' problem now is to maximize r_e , as given by (4), subject to the constraint (5).

The solution is apparent if equation (4) is rearranged:

$$(6) \quad r_e = \frac{1}{f_e} \left[\frac{pq - wL}{K} \right] - \frac{i_d f_d}{f_e}.$$

By equation (6), r_e increases with an increase in the bracketed term (rate of return to total capital), which is just the left-hand side of (5). Hence, the managers' solution is to operate the firm with return on total capital at its maximum allowed value, f .

To obtain a geometric interpretation, it is again convenient to rearrange the ex-

³ A composite management objective function is considered in Bailey and Malone.

⁴ Financial strategy with respect to debt ratio is an involved issue, with opinions ranging from the Miller-Modigliani thesis that debt ratio has negligible effect on stockholder behavior to the belief that there is an optimum value of f_d . This controversy will not be dealt with here; it is simply observed that it is common for a public utility's debt ratio to be relatively constant over time.

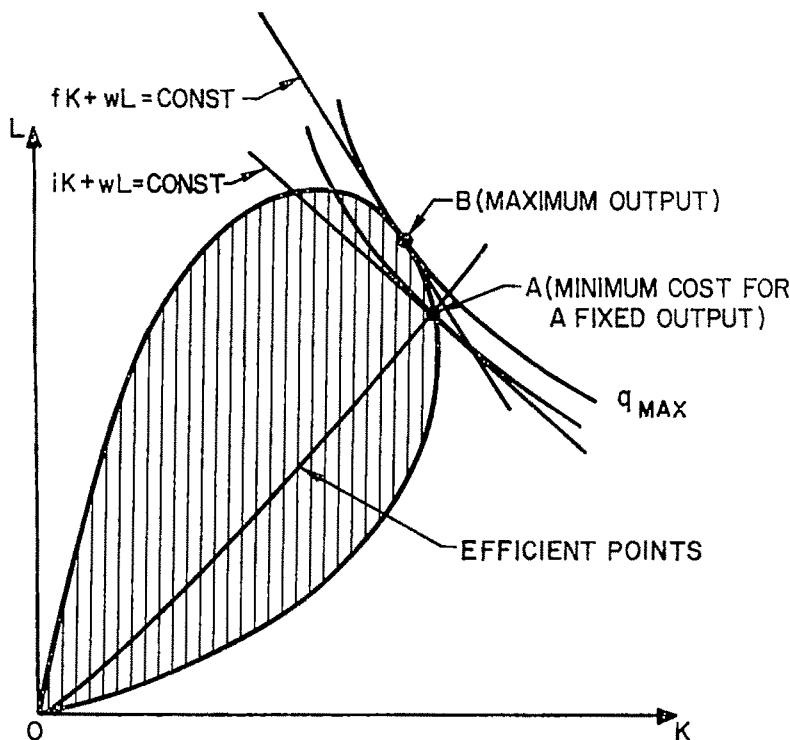


FIGURE 4

Solution for the Case of Maximum Rate of Return to Equity.

pression for the rate of return to total capital $r = (pq - wL)/K$ into the form:

$$\pi = (r - i)K.$$

The locus of K, L values corresponding to both constant r_s and r is thus again given by the intersection of a plane hinged on the L -axis and the profit hill. At the maximum allowable rate to return to equity (or to total capital), $r = f$ and this plane simply corresponds to the constraint plane.

However, it is important to note that in contrast to the profit maximizer, the rate of return maximizer is indifferent as to where he operates along the intersection of the plane $\pi = (r - i)K$ and the profit hill (or, in terms of Figure 4, as to where he operates along the constraint curve). All (K, L) input combinations which yield the maximum rate of return allowed by law are equally attractive.

Faced with this indifference, one looks for other criteria which will lead the firm to pick unique K and L values. Two plausible criteria are the following:

- (1) The firm will try to operate at a point where cost is minimized for the amount of output. In this case, the operating point must be an efficient point. Since the firm also desires to operate on the constraint curve, at maximum allowed r_s , the operating point will be at the intersection of the locus of efficient points and the constraint curve (point A, Figure 4).
- (2) The firm will try to maximize output. It will then operate at point B in Figure 4, on the isoquant which is tangent to the constraint curve. It is easily shown that at point B the constraint curve and the isoquant are both tangent to a level line of $fK + wL$.

With $f > i$, level lines of $fK + wL$ are steeper than those of $iK + wL$ (Figure 4). Thus, at B , the firm is undercapitalized for its production level, and the effect is the opposite of the AJ effect.

IV. Labor Constraint

The modern utility operates under constraints other than rate of return regulation, e.g., regulation on quality of service and maintenance of safety standards. Dealings with labor are a source of important additional constraints, although these are less direct than regulation and more sporadic. A simple model of this constraint is the following: Assume the firm's labor force will strike if the money available for wages, $pq - iK$, rises above a threshold value. That is, the firm could pay a rate of ψ , given by $pq - iK = \psi L$; a strike will occur if ψ rises above a value ψ^* . Hence the firm is bound by

$$pq - iK \leq \psi^* L.$$

Likewise, it is pressured to keep wages at w to avoid inflation. Subject to these constraints, it is assumed that the firm wishes to maximize profit, π .

This model is exactly the original AJ model, but with the roles of capital and labor reversed. Again, leaving dynamic factors aside, the firm's solution is obvious. It operates at the point of maximum labor input along the constraint curve formed by the intersection of the profit hill and the plane $\pi = (\psi^* - w)L$. In this case, it operates inefficiently with overintensive labor rather than capital use.

V. Addition of Nonproductive Plant

An ancient allegation in regulation is that regulated firms will tend to "pad the rate base," by adding useless plant in order to increase allowable earnings. The allegation is usually countered by the

argument that the firm has little incentive to acquire useless capital when, for the same rental, it can obtain productive capital.

The issue can be examined from the point of view of the theory of the firm as follows. Assume that an amount of nonproductive capital K^* is added to the plant to give a total capital of $K + K^*$.⁵ The production function, $q = q(K, L)$ now involves only productive capital, and profit is

$$(7) \quad \pi = pq - wL - i(K + K^*),$$

since the firm must pay for K^* even though it is useless. The regulatory constraint becomes

$$(8) \quad \pi \leq (f - i)(K + K^*).$$

From (7) it is seen that the cost of the useless capital K^* lowers the entire profit hill by the amount iK^* , while by (8), the constraint plane is raised parallel to itself so that it cuts the K -axis at $-K^*$. Geometrically, the situation is that shown in Figure 5. Comparison with Figure 1 shows that raising of the constraint plane is counteracted by the lowering of the profit hill, which corresponds to the classical allegation and counterargument.

However, the addition of useless capital cannot result in the profit maximum of Figure 5 being higher than that of Figure 1. For suppose the firm were operating at a constrained maximum with useless capital K^* as in Figure 5. It could exchange K^* for an equal amount of productive capital and at the same time reduce the labor factor so as to remain on the same isoquant. In this move, the firm's revenue and rate base would remain the same, but its labor expenses would be reduced, resulting

⁵ The idea of dividing total capital into useful and nonproductive portions for purposes of the modelling of rate base padding was first suggested by G. R. Faulhaber in an unpublished note.

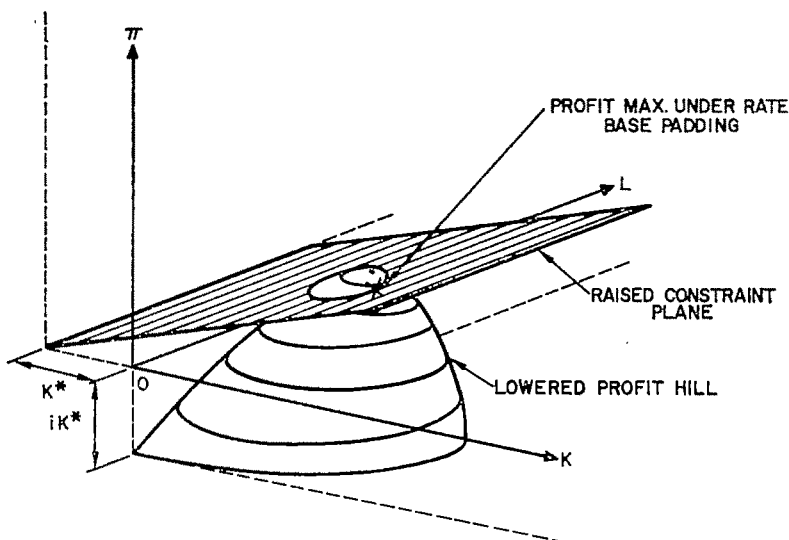


FIGURE 5
Geometric Representation of Rate Base Padding

in a profit increase and hence violation of the constraint. This would put the firm somewhere in the prohibited shaded region of Figures 2-4. It could then further move along the isoquant to the constraint curve by adding more productive capital and retiring more labor. But this move would expand its rate base and hence its profit would be higher than at the original, useless-capital constrained maximum. Once on the constraint curve, the firm could realize a still higher constrained profit by moving to the K_{\max} point. Hence, the constrained profit maximum of Figure 1, using only productive capital, is always higher than the constrained useless-capital profit maximum of Figure 5.

REFERENCES

- H. Averch and L. L. Johnson, "Behavior of the Firm Under Regulatory Constraints," *Amer. Econ. Rev.*, Dec. 1962, 52, 1053-69.
- E. E. Bailey and J. C. Malone, "Resource Allocation and the Regulated Firm," *Bell J. Econ. Manag. Sci.*, forthcoming.
- A. E. Kahn, "The Graduated Fair Return: Comment," *Amer. Econ. Rev.*, Mar. 1968, 58, 170-3.
- A. K. Klevorick, "The Graduated Fair Return: A Regulatory Proposal," *Amer. Econ. Rev.*, June 1966, 56, 477-84.
- H. M. Trebing, ed., *Performance Under Regulation*, East Lansing 1968.
- F. M. Westfield, "Regulation and Conspiracy," *Amer. Econ. Rev.*, June 1965, 55, 424-43.

Migration, Unemployment and Development: A Two-Sector Analysis

By JOHN R. HARRIS AND MICHAEL P. TODARO*

Throughout many less developed economies of the world, especially those of tropical Africa, a curious economic phenomenon is presently taking place. Despite the existence of positive marginal products in agriculture and significant levels of urban unemployment, rural-urban labor migration not only continues to exist, but indeed, appears to be accelerating. Conventional economic models with their singular dependence on the achievement of a full employment equilibrium through appropriate wage and price adjustments are hard put to provide rational behavioral explanations for these sizable and growing levels of urban unemployment in the absence of absolute labor redundancy in the economy as a whole. Moreover, this lack of an adequate analytical model to account for the unemployment phenomenon often leads to rather amorphous explanations such as the "bright lights" of the city acting as a magnet to lure peasants into urban areas.

In this paper we shall diverge from the usual full employment, flexible wage-price models of economic analysis by formulating a two-sector model of rural-urban migration which, among other things, recognizes the existence of a politically

determined minimum urban wage at levels substantially higher than agricultural earnings.¹ We shall then consider the effect of this parametric urban wage on the rural individual's economic behavior when the assumption of no agricultural labor surplus is made, i.e., that the agricultural marginal product is always positive and inversely related to the size of the rural labor force.² The distinguishing feature of this model is that migration proceeds in response to urban-rural differences in *expected earnings* (defined below) with the urban employment rate acting as an equilibrating force on such migration.³ We shall then use the overall model for the following purposes:

- 1) to demonstrate that given this po-

¹ For some empirical evidence on the magnitude of these real earnings differentials in less developed economies, see Reynolds, Berg, Henderson, and Ghai.

² We do not make the special assumption of an agricultural labor surplus for the following reasons: Most available empirical evidence to date tends to cast doubt on the labor surplus argument in the context of those economies of Southeast Asia and Latin America where such a surplus would be most likely to exist (see Kao, Anschel, and Eicher). Moreover, few if any economists would seriously argue that general labor surplus exists in tropical Africa, the area to which this paper is most directly related.

³ For a dynamic model of labor migration in which urban unemployment rates and expected incomes play a pivotal role in the migration process, see Todaro. However, unlike the present model which attempts to view the migration process in context of aggregate and intersectoral welfare considerations, Todaro's model was strictly concerned with the formulation of a positive theory of urban unemployment in developing nations. As such, it did not specifically consider the welfare of the rural sector, nor was it concerned with the broader issues of economic policy considered in the present paper.

* The authors are assistant professor of economics, Massachusetts Institute of Technology and research fellow, Institute for Development Studies, University College, Nairobi, respectively. They would like to thank the Rockefeller Foundation for making possible their research on economic problems of East Africa. Peter Diamond, Richard Eckaus, Joseph Stiglitz, two anonymous referees, and the managing editor made valuable comments on a previous draft. The authors, of course, are responsible for remaining errors.

litically determined high minimum wage, the continued existence of rural-urban migration in spite of substantial overt urban unemployment represents an economically rational choice on the part of the individual migrant;

2) to show that economists' standard policy prescription of generating urban employment opportunities through the use of "shadow prices" implemented by means of wage subsidies or direct government hiring will *not* necessarily lead to a welfare improvement and may, in fact, exacerbate the problem of urban unemployment;

3) to evaluate the welfare implications of alternative policies associated with various back-to-the-land programs when it is recognized that the standard remedy suggested by economic theory—namely, full wage flexibility—is for all practical purposes politically infeasible. Special attention will be given here to the impact of migration cum unemployment on the welfare of the rural sector as a whole which gives rise to intersectoral compensation requirements; and, finally,

4) to argue that in the absence of wage flexibility, an optimal policy is, in fact, a "policy package" including *both* partial wage subsidies (or direct government employment) and measures to restrict free migration.

I. *The Basic Model*

The basic model which we shall employ can be described as a two-sector internal trade model with unemployment. The two sectors are the permanent urban and the rural. For analytical purposes we shall distinguish between sectors from the point of view of production and income. The urban sector specializes in the production of a manufactured good, part of which is exported to the rural sector in exchange for agricultural goods. The rural sector has a choice of either using all available labor to produce a single agricultural good, some

of which is exported to the urban sector, *or* using only part of its labor to produce this good while *exporting* the remaining labor to the urban sector in return for wages paid in the form of the manufactured good. We are thus assuming that the typical migrant retains his ties to the rural sector and, therefore, the income that he earns as an urban worker will be considered, from the standpoint of sectoral welfare, as accruing to the rural sector.⁴ However, this assumption is not at all necessary for our demonstration of the rationality of migration in the face of significant urban unemployment.

The crucial assumption to be made in our model is that rural-urban migration will continue so long as the *expected* urban real income at the margin exceeds real agricultural product—i.e., prospective rural migrants behave as maximizers of *expected* utility. For analytical purposes, we shall assume that the total urban labor force consists of a permanent urban proletariat without ties to the rural sector plus the available supply of rural migrants. From this combined pool of urban labor, we assume that a periodic *random job selection process* exists whenever the number of available jobs is exceeded by the number of job seekers.⁵ Consequently, the expected

⁴ In tropical Africa especially, this notion that migrants retain their ties to the rural sector is quite common and manifested by the phenomenon of the extended family system and the flow of remittances to rural relatives of large proportions of urban earnings. However, the reverse flow, i.e., rural-urban monetary transfers is also quite common in cases where the migrant is temporarily unemployed and, therefore, must be supported by rural relatives. For an excellent discussion of this phenomenon from a sociological point of view, see Gugler (pp. 475–78).

⁵ The qualitative conclusions of the model do not depend on the precise nature of the selection process. We have assumed random selection not merely for analytic convenience but also because it directly corresponds to an appropriate dynamic construct developed in Todaro's 1969 article. There it is shown that over time expected and actual earnings will converge to a positive number even though the rate of job creation is less than the rate of migration so that unemployment is increasing.

urban wage will be defined as equal to the fixed minimum wage (expressed in terms of manufactured goods) times the proportion of the urban labor force actually employed (see equation (6)). Finally, we assume perfectly competitive behavior on the part of producers in both sectors with the further simplifying assumption that the price of the agricultural good (defined in terms of manufactured goods) is determined directly by the relative quantities of the two goods produced.

Consider now the following formulation of the model.

Agricultural Production Function:

$$(1) \quad X_A = q(N_A, \bar{L}, \bar{K}_A), \quad q' > 0, \quad q'' < 0$$

where,

X_A is output of the agricultural good,
 N_A is the rural labor used to produce this output,

\bar{L} is the fixed availability of land,

\bar{K}_A is the fixed capital stock,

q' is the derivative of q with respect of N_A , its only variable factor.

Manufacturing Production Function:

$$(2) \quad X_M = f(N_M, \bar{K}_M), \quad f' > 0, \quad f'' < 0$$

where

X_M is the output of the manufactured good,

N_M is the total labor (urban and rural migrant) required to produce this output.

\bar{K}_M is fixed capital stock, and

f' is the derivative of f with respect to N_M , its only variable factor.

It is interesting to note in this context that sociologist Gugler who has spent considerable time studying labor migration in Africa has recently concluded that rural-urban migration is essentially an economic phenomenon that can be portrayed as a "game of lottery" in which rural migrants come to the city fully aware that their chances of finding a job are low. However, the great disparity between urban and rural wages makes the successful location of an urban salaried job so attractive that unskilled migrants are willing to take a chance (pp. 472-73). See also Hutton.

Price Determination:

$$(3) \quad P = \rho \left(\frac{X_M}{X_A} \right), \quad \rho' > 0$$

where

P , the price of the agricultural good in terms of the manufactured good, (i.e., the terms of trade) is a function of the relative outputs of agricultural and manufactured good when the latter serves as numeraire.⁶

Agricultural Real Wage Determination:

$$(4) \quad W_A = P \cdot q'$$

where

W_A , the agricultural real wage, is equal to the value of labor's marginal product in agriculture expressed in terms of the manufactured good.

Manufacturing Real Wage:

$$(5) \quad W_M = f' \geq \bar{W}_M.$$

The real wage in manufacturing, expressed in terms of manufactured goods, is equated with the marginal product of labor in manufacturing because of profit maximization on the part of perfectly competitive producers. However, this wage is constrained to be greater than or equal to the fixed minimum urban wage. In our analysis, we shall be dealing only with cases in which $f' = \bar{W}_M$ (i.e., there is never an excess demand for labor at the minimum wage).

Urban Expected Wage:

$$(6) \quad W_u^* = \frac{\bar{W}_M N_M}{N_u}, \quad \frac{N_M}{N_u} \leq 1,$$

⁶ A sufficient, but not necessary, condition for this assumption is that all individuals in the economy have the same homothetic preference map. Again, the assumption is made for analytical convenience. The qualitative conclusions of our analysis will remain unaffected under several plausible assumptions about distribution of income and tastes.

where the *expected* real wage in the urban sector, W_u^e , is equal to the real minimum wage \bar{W}_M adjusted for the proportion of the total urban labor force (permanent urban plus migrants, denoted as N_u) actually employed, N_M/N_u .⁷ Only in the case of full employment in the urban sector ($N_M = N_u$) is the expected wage equal to the minimum wage (i.e., $W_u^e = \bar{W}_M$).

Labor Endowment:

$$(7) \quad N_A + N_u = \bar{N}_R + \bar{N}_u = \bar{N}$$

There is a *labor constraint* which states that the sum of workers actually employed in the agricultural sector (N_A) plus the total urban labor force (N_u) must equal the sum of initial endowments of rural (\bar{N}_R) and permanent urban (\bar{N}_u) labor which in turn equals the total labor endowment (\bar{N}).

Equilibrium Condition:

$$(8) \quad W_A = W_u^e$$

Equation (8), an equilibrium condition, is derived from the hypothesis that migration to the urban area is a positive function of the urban-rural *expected* wage differential. This can be written formally as

$$(9) \quad \dot{N}_u = \psi \left(\frac{\bar{W}_M N_M}{N_u} - P \cdot q' \right),$$

$$\psi' > 0, \quad \psi(0) = 0$$

where \dot{N}_u is a time derivative. Clearly then, migration will cease only when the expected income differential is zero, the con-

dition posited in (8).⁸ It is important to note that this assumes that a migrant gives up only his marginal product.⁹

We thus have 8 equations in 8 unknowns X_A , X_M , N_A , N_M , W_A , W_u^e , N_u and P . Given the production functions and fixed minimum wage \bar{W}_M , it is possible to solve for sectoral employment, the equilibrium unemployment rate and, consequently, the equilibrium expected wage, relative output levels and terms of trade. Let us analyze how such an unemployment equilibrium can come about.

The essence of our argument is that in many developing nations the existence of an institutionally determined urban minimum wage at levels substantially higher than that which the free market would allow can, and usually does, lead to an equilibrium with considerable urban unemployment. In our model migration is a disequilibrium phenomenon. In equilibrium $\bar{W}_M N_M / N_u = Pq'$ and migration ceases. (See Appendix I for proof that this equilibrium is stable.) Now we know from equation (5) that in the competitive urban manufacturing sector, $\bar{W}_M = f'$. We also know from equation (7) that $\bar{N} - N_A = N_u$ and from equation (3) that $P = \rho(X_M/$

⁸ $\psi(0) = 0$ is purely arbitrary. If, instead, we assume $\psi(\alpha) = 0$ where α can take on any value, migration will cease when the urban-rural expected wage differential is equal to α . None of the subsequent analysis is affected qualitatively by specifying $\alpha = 0$. Equation (8) would merely be written as $W_A + \alpha = W_u^e$.

⁹ Other assumptions could be made. Much of the literature has stressed that in peasant economies producers receive their average product which is higher than their marginal product. Indeed, this is at the heart of the well-known Lewis and Fei-Ranis models. However, these models ignore the migration decision and seem to assume that migrants continue to receive their share of peasant production yet migrate only if jobs are actually available. In much of Africa it appears that migrants continue to receive income from land after migration and commonly hire labor to work on their farms in their absence. There is also a considerable group of landless individuals who work on farms for wages. Thus it would appear that our assumption is not unreasonable. The analysis could easily be modified to make earnings foregone equal to average product, however.

⁷ This assumes a very particular form of wage expectation, namely that the expected wage is equal to the average urban wage. Although this is a convenient expression to work with, we could be more general and make the expected wage some function of the average urban wage. Indeed, the only restrictions on such a function that are necessary for our results are that, *ceteris paribus*, the expected wage varies directly with the minimum wage and inversely with the unemployment rate.

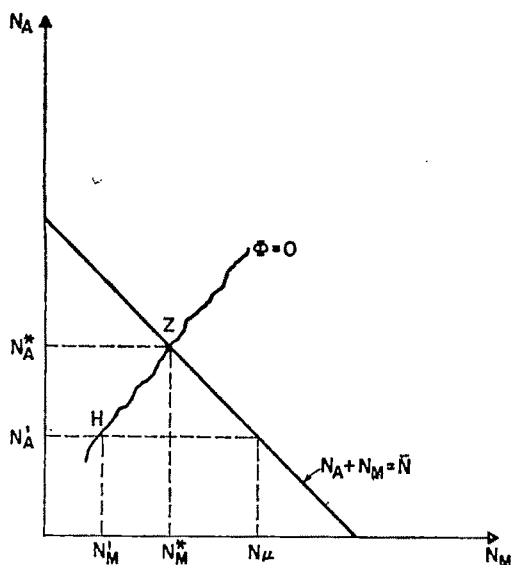


FIG 1.

X_A). Therefore, we can rewrite our equilibrium condition (8) as

$$(8') \quad \Phi = \rho(X_M/X_A)q' - \frac{f'N_M}{\bar{N} - N_A} = 0.$$

Since X_M and X_A are functions of N_M and N_A respectively, Φ is an implicit function in N_A and N_M which, for any stated minimum wage, can be solved for the equilibrium combination of agricultural and manufacturing employment. From this solution the levels of urban unemployment and commodity outputs can also be determined. There will be a unique equilibrium associated with each possible value of the minimum wage, and the locus of these equilibria is plotted in Figure 1 as the line $\Phi=0$ in N_A, N_M space.¹⁰ The line N_A

¹⁰ In Figure 1 we have assumed that

$$\frac{dN_A}{dN_M} = -[\Phi_{N_M}/\Phi_{N_A}] > 0$$

although this need not necessarily hold true. Differentiating (8') partially with respect to N_A we find that

$$\Phi_{N_A} = \frac{-\rho'f'q'^2}{q^3} + \rho q'' - \frac{\rho q'}{\bar{N} - N_A}$$

$+N_M = \bar{N}$ in Figure 1 is the locus of full-employment points.

Point Z is the only equilibrium full-employment point in Figure 1 at which N_M^* workers would be employed in manufacturing and N_A^* in agriculture. Points on the locus $\Phi=0$ east of Z are infeasible and will not be considered further, while points to the west of Z are associated with min-

which is unambiguously negative since $q'' < 0$ and $\rho' > 0$. Differentiating (8') partially with respect to N_M we find that

$$\Phi_{N_M} = \frac{1}{\eta_{LW}} - \eta_P \frac{f'N_M}{X_M} + 1$$

which is less than, equal to, or greater than zero as

$$-\frac{1}{\eta_{LW}} + \eta_P \frac{f'N_M}{X_M} \begin{matrix} > \\ = \\ < \end{matrix} 1,$$

where

$$\eta_{LW} = -\frac{dN_M}{dW} \cdot \frac{W}{N_M}$$

is the wage elasticity of demand for labor and

$$\eta_P = \frac{dP}{d\left(\frac{X_M}{X_A}\right)} \cdot \frac{X_M/X_A}{P}$$

is the elasticity of the terms of trade with respect to a change in relative outputs. It follows, therefore that the slope of the locus of equilibria, dN_A/dN_M depends on the respective employment and price elasticities.

A sufficient condition for Φ_{N_M} to be negative (making dN_A/dN_M positive) is for the wage elasticity of employment to be less than one, a situation which recent empirical studies suggest is likely to exist (see Erickson, Harris and Todaro (1969), and Katz). However, even if η_{LW} exceeds unity, dN_A/dN_M can still be positive providing price elasticity is sufficiently high. The logic of these conditions is clear. If η_{LW} is less than one, a decline in the minimum wage will lower the urban wage bill even though employment and output increase. This causes the expected urban wage to decline thereby reducing the expected rural-urban earnings differential which gives rise to reverse migration and increased rural employment and output. If η_{LW} exceeds unity, a fall in the minimum wage is accompanied by an increased urban wage bill and, hence, a higher expected urban wage. However, the expected rural-urban earnings differential can either increase or decrease in this case depending on the movement in terms of trade which raises the value of the marginal product in agriculture. For example, if η_{LW} were 1.5 and the wage share of manufacturing output ($f'N_M/X_M$) were .50, then an agricultural price elasticity greater than 0.67 would be sufficient to make dN_A/dN_M positive.

imum wages higher than the full-employment wage. There is a monotonic mapping such that higher minimum wages are associated with points on $\Phi=0$ lying farther to the west. Thus we can demonstrate that the setting of a minimum wage above the market-clearing level causes an economy to settle at a point such as H in Figure 1. At H , N'_A workers are employed in agriculture, N'_M in manufacturing, and $N_u - N'_M$ workers are unemployed. It is evident that the minimum wage causes a loss of employment and hence output in both sectors.¹¹

It is important to note that even though an equilibrium at point H represents a suboptimum situation for the economy as a whole, it does represent a rational, utility maximizing choice for individual rural migrants given the level of the minimum wage.

One final point might be raised at this juncture. So far we have assumed that the urban minimum wage is fixed in terms of the manufactured good. What if, instead, the minimum wage were fixed in terms of the agricultural good? We would then substitute for equation (5):

$$(5') \quad W_M = \frac{f'}{P} \geq \bar{W}_M.$$

Substituting (4), (5'), and (6) into (8) we get the equilibrium relationship

$$(11) \quad Pq' = \frac{\left(\frac{\bar{f}'}{P}\right) \cdot N_M}{N_u}.$$

¹¹ If $dN_A/dN_M < 0$, which we believe to be empirically unlikely, this statement would have to be modified. In such a case, increasing the minimum wage will decrease manufacturing employment but will increase agricultural employment and output. Unemployment will result from the imposition of a minimum wage but we can no longer assert that the level of unemployment will increase concomitantly with the level of the minimum wage.

We can then imagine an economy starting initially at the point on the production possibilities frontier at which X_M is that for which equation (5') is satisfied and assume that

$$Pq' < \frac{\left(\frac{\bar{f}'}{P}\right) \cdot N_M}{N_u}$$

at that point. The equilibrium point will again be reached through a simultaneous raising of Pq' and lowering of W_u^* in response to migration. As relative agricultural output falls, P will rise. This in turn will cause output of the manufactured good to fall as well, since producers will produce up to the point that $f' = \bar{W}_M P$ which rises in terms of the manufactured good. Note that f' can be raised only through output restriction (since $f'' < 0$). Therefore, in general, we would find that imposition of a minimum wage gives rise to an equilibrium characterized by unemployment and loss of potential output of both goods. A new locus $\Phi'=0$ will be defined in Figure 1 such that the point on Φ' corresponding to any given minimum wage will be west of the corresponding point on Φ .

Although our initial assumption is a bit easier to handle, the principal conclusion remains unaffected if we make the minimum wage fixed in terms of the agricultural good. Equilibrium is only achievable with unemployment. Actual minimum wage setting is usually done with reference to some general cost of living index, and food is the largest single item in the budget of most urban workers. (See Massell and Heyer, and the Nigeria report.) Hence, the second case may be somewhat more realistic. Note that in the first case the "true" real wage was reduced somewhat by the rising agricultural price, while in the latter case it is increased by the falling relative price of the manufactured good.

III. *Implications for Development Policy*

A. *Planning in Terms of Shadow Prices*

The standard solution to the problem of an institutionally determined wage that is higher than the equilibrium level is to employ labor in the public sector according to a shadow wage and/or to grant a payroll subsidy to private employers that equates private costs with this shadow wage.¹² Two main problems arise with this prescription: first, how can one determine the appropriate shadow wage? and, secondly, what are the implications of executing such a scheme when the institutional wage will continue to be paid to the employed? Our model can shed light on both of these issues.

In a static framework the appropriate shadow wage is the opportunity cost of labor hired by the industrial sector. Hence, if labor is hired to the point that its marginal product in industry is equated with the shadow wage which in turn is equated with the marginal product in agriculture, marginal productivity of labor will be equal in both sectors, a necessary condition for an optimal allocation of resources. Naturally, this assumes a positive marginal product in agriculture and sufficient factor mobility to ensure full employment of labor. The existence of urban unemployment, however, suggests that there may be

a pool of labor that can be tapped without sacrificing output. Consequently, it might be suggested that even though agricultural labor is fully employed at peak seasons, the appropriate shadow wage for urban labor is likely to be one that is lower than the marginal product in agriculture. This would be correct if the two labor forces, urban and rural, were separate noncompeting groups. In linear programming terms, there are two labor constraints and each may well have a different associated shadow wage.

Now, the essence of our model is that the two sectors *are* intimately connected through labor migration. If one additional job is created in the industrial sector at the minimum wage, the expected wage will rise and rural-urban migration will be induced. In Appendix II it is shown that more than one agricultural worker will likely migrate in response to the creation of one additional industrial job. Hence, the opportunity cost of an industrial worker will exceed the marginal product of an agricultural worker. On the other hand, an increase in agricultural income will induce reverse migration with no diminution of industrial output. Thus, the opportunity cost of labor is lower to the agricultural than to the industrial sector!

The literature has been strangely silent for the most part about the full implications of using shadow-wage criteria. In a static context, Stolper has pointed out that financing subsidies or losses of public enterprises gives rise to fiscal problems, but unfortunately this issue has not yet been pursued in sufficient detail.¹³ If the problem is considered at all, the analyst usually assumes that a system of nondistorting lump-sum taxes is available. Little, Lefebvre, and

¹² Hagen (p. 498) states, "a subsidy per unit of labor equal to the wage differential [between agriculture and industry] will increase real income further [than a tariff] and if combined with free trade will permit attaining an *optimum optimorum*." Bardhan (p. 379) similarly adds, "The best remedy for the misallocation caused by a wage differential is . . . an appropriate subsidy to the use of labor in the manufacturing industry." It is important to recall that this argument is dependent on variable proportions production functions. If production coefficients are fixed, a wage subsidy will have no effect in the short run. The classic statement of this case is by Eckaus. Bardhan explores its implications for subsidy in a dynamic context. Both of these papers, however, posit surplus labor in agriculture, an assumption we do not wish to make in an African context.

¹³ Lefebvre assumes that a wage subsidy can be financed by a profits tax, while other writers, e.g. Hagen, Bardhan, and Chakravarty never even consider the problem. Even Little and Mirrlees who present an excellent discussion of how to calculate a shadow wage never mention the fiscal problems of implementation.

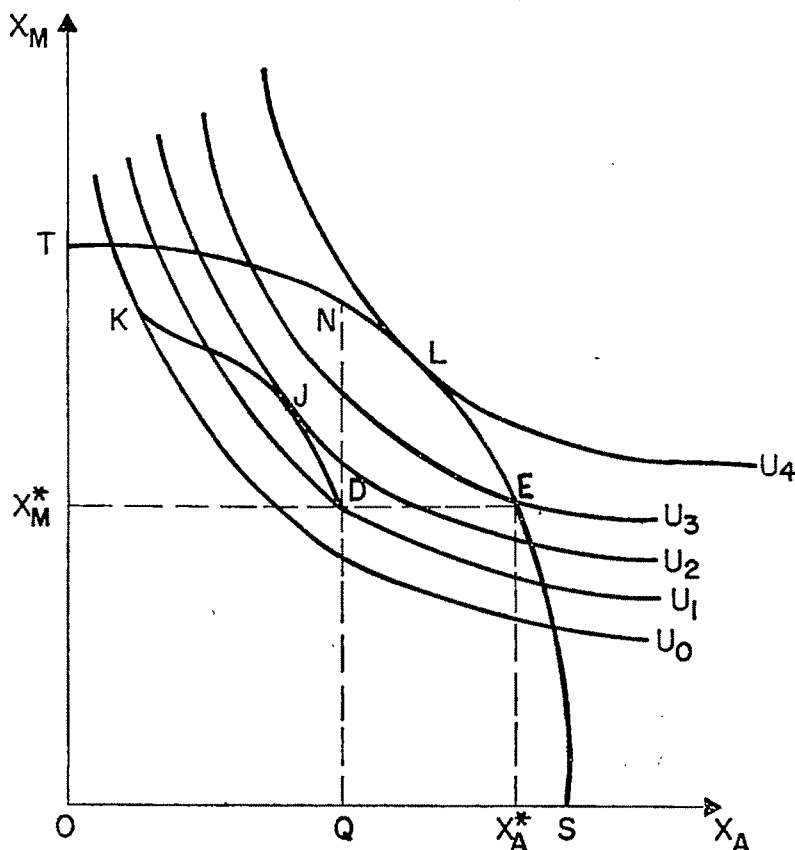


FIGURE 2

Little and Mirrlees have pointed out that in a dynamic setting, the extra consumption arising from payment of the institutional wage diverts resources from investment to consumption; thus some of the foregone future consumption should be considered in calculating the shadow wage. In our model, payment of the minimum wage to additional industrial workers will induce more rural-urban migration. Therefore, implementation of a shadow-wage employment criterion will have important effects on the level of agricultural output and on urban unemployment. The argument can be clarified with reference to Figure 2.

The initial equilibrium, given the minimum wage, is at point *D* with output of

the manufactured good restricted to OX_M^* . If individuals did not migrate in response to expected wage differentials, the economy could produce at point *E*, but migration reduces agricultural output to the level OQ . The theory of shadow pricing suggests that with an appropriate wage subsidy (or public-sector-hiring rule) the economy could move to point *L* on the production possibilities frontier which, with the posited social indifference map, is the optimum position. Welfare would be increased from a level U_1 to a higher level U_4 .

In the context of our model, such a point is unattainable. The effect of implementing a shadow wage will be to increase production of the manufactured good. But creation of an additional job at the minimum

wage will induce some additional migration (see Appendix II) from the rural sector and therefore agricultural output will fall. Hence, movement from D can only be in a northwest direction. The line DK in Figure 2 is the locus of all such attainable points and it is evident that there is only one point, K , at which there can be full employment of the economy's labor resources. At that point the expected wage will be equal to the minimum wage since there is no urban unemployment. Therefore, the marginal product in agriculture will have to be equal to the minimum wage. But, with the subsidy, the marginal product of labor in manufacturing will be lower than in agriculture, hence K lies inside the production possibilities frontier. (In the extreme case in which marginal productivity in agriculture can never be as high as the minimum wage, K will coincide with T , the point of complete specialization in manufactures.) This situation will certainly not meet the conditions for a general optimum which can be met only at L . Thus, implementing a shadow wage criterion to the point that urban unemployment is eliminated will not generally be a desirable policy.¹⁴

However, some level of wage subsidy will usually lead to an improvement. In Figure 2 it is clear that point J , with a welfare level U_2 , will be preferable to D . The criterion for welfare maximization, derived in Appendix III, is the following:

$$(12) \quad f' = Pq' \left(\frac{dN_u}{dN_M} \right).$$

Note what this means. Creating one additional job in the industrial sector increases output by f' but, since increased

employment will raise the expected urban wage, migration will be induced in an amount dN_u/dN_M . The right-hand side of equation (12) states the amount of agricultural output sacrificed because of migration. Thus the shadow wage will be equal to this opportunity cost of an urban job and the amount of subsidy will be $\bar{W}_M - f'$. So long as $f' > Pq' (dN_u/dN_M)$, aggregate welfare can be increased by expanding industrial employment through subsidy or public sector hiring. Clearly the more responsive is migration to industrial employment, the higher is the social cost of industrialization and the smaller is the optimal amount of subsidy. In many African economies it is likely that dN_u/dN_M exceeds unity. If so, it will be optimal for the marginal product of labor in industry to be higher than in agriculture and urban unemployment will be a persistent phenomenon so long as minimum wages are set above a market-clearing level.

The discussion so far has ignored two other adverse effects of using a shadow wage. As mentioned earlier, several writers have noted that payment of a subsidized minimum wage to additional workers will increase total consumption, thereby reducing the level of resources available for investment. If foregone future consumption is positively valued, the opportunity cost of industrial labor will be higher than indicated in equation (12) and the shadow wage will be raised correspondingly. Furthermore, wage subsidies or public enterprise losses must be financed and if revenue cannot be raised through costless lump-sum taxes, the opportunity cost of raising taxes must be considered. Both of these effects will reduce the desirable amount of subsidized job creation in the industrial sector.

It is interesting to note that this model implies different opportunity costs of labor to the two sectors. While the creation of an additional job in the urban area reduces

¹⁴ As shown in Appendix III, DK is not uniformly convex. Therefore, K may be the best attainable point in some cases and the first-order conditions may not ensure optimality. As drawn in Figure 2, moving from D to K represents a worsening of welfare, but this clearly is not a necessary conclusion.

agricultural output through induced migration, additional employment can be generated in the agricultural sector without reducing manufacturing output. If this phenomenon is not taken into account, standard application of investment criteria is likely to be biased in favor of urban projects.

B. *Migration Restriction*

An alternative approach to the problem of urban unemployment is to physically control migration from the rural areas. Such controls have recently been introduced in Tanzania and have been used for some time in South Africa.¹⁵ Other countries, such as Kenya, are giving serious consideration to instituting such a policy. Although we personally have grave reservations about the ethical issues involved in such a restriction of individual choice and the complexity and arbitrariness of administration, it seems desirable to investigate the economic implications of such a policy.

Looking at Figure 2 it is obvious that with the minimum wage such that industrial output is OX_M^* , prohibition of migration in excess of the labor required to produce that output will allow the economy to produce at point E . The movement from D to E arising from restriction of migration leads to an unambiguous aggregate welfare improvement providing appropriate lump-sum redistribution is effected. Since such compensation is notoriously difficult to carry out in practice, it will be useful to examine the welfare implications of such a move on each of the two sectors in the absence of compensation.

Recall that the two sectors were defined to be a permanent urban group and a rural sector that produces both agricultural goods and exports labor to the urban area

in exchange for wages in the form of manufactured goods.¹⁶ In Figure 3 the line $T'S'$ represents production possibilities for the agricultural sector when labor export is allowed. If its entire labor endowment is devoted to agricultural production, it can produce a quantity OS' . However, by exporting its labor, the agricultural sector can "produce" the manufactured good (wages are paid in the form of this good). Hence this production possibilities frontier depends on market forces (wage levels and unemployment) as well as on purely technological factors. The amount of agricultural output foregone if a unit of labor is to be "exported" is its marginal product; the amount of manufactured goods obtained by the exported labor unit depends on the wage, the amount of employment obtained by the exported unit, and its effect on employment of previously exported units.

In addition to these production possibilities, the rural sector also has the opportunity to trade some of its agricultural output with the permanent urban sector in exchange for manufactured goods. Corresponding to each point on the production possibilities frontier $T'S'$, there is a determinate price of the agricultural good. The manner in which alternative constellations of production and trade affect the

¹⁵ See Harris and Todaro (1969) for an analysis of the Tanzanian program.

¹⁶ In considering the welfare of the rural sector as a whole we are making the tacit assumption that there is redistribution of goods between individuals in this sector. This is a very strong assumption. Yet there is considerable evidence from tropical Africa that employed urban migrants repatriate substantial portions of their earnings to their kinsmen remaining in the rural areas and conversely that income both in cash and kind is received by unemployed migrants from kinsmen remaining on the farm. To the extent that the extended family system does redistribute goods between members, this assumption may be tenable as a first approximation. As Gugler (p. 480) has pointed out, it is appropriate to view the extended family as maximizing its income by allocating its members between agriculture and urban wage employment. Although there is some evidence that growing numbers of urban workers are settling permanently and gradually eliminating rural ties, it will be many years before such ties are completely severed.

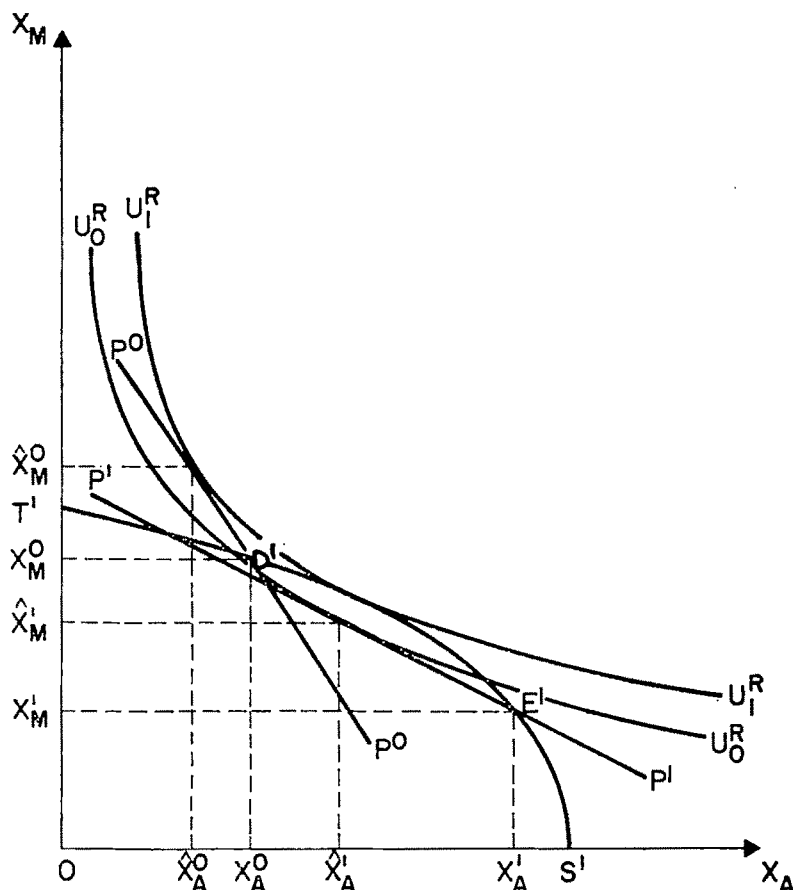


FIGURE 3

sector's welfare can be illustrated by Figure 3.

D' corresponds to the initial unemployment equilibrium D (Figure 2). At that point the rural sector as a whole "produces" X_A^0 and X_M^0 of the two goods. It also has the opportunity to trade at the price P^0 . By trading some of its agricultural output to the permanent urban sector for additional manufactured goods, it consumes \hat{X}_A^0 , \hat{X}_M^0 and achieves a welfare level of U_I^R . Restriction of migration results in the sector's producing X_A^1 , X_M^1 . If it could still trade at price P^0 , the agricultural sector would clearly be better off. But this is impossible. At E' (which corresponds to E in Figure 2), the price of

agricultural good will fall to P' and with trade the best consumption bundle attainable by the sector is \hat{X}_A , \hat{X}_M which corresponds to a lower level of welfare U_O^R . (Note that if P' did not cut $T'S'$ there could be no incentive to migrate at E' .)

It can be shown that $Pq'(1 - 1/\eta)$ (where η is the price elasticity of demand for the agricultural good) is the amount of the manufactured good sacrificed by the rural sector as a result of removing one worker from producing the agricultural good which could have been exchanged for the manufactured good at the market price $1/P$. This quantity is less than the value of labor's marginal product in agriculture (Pq') since the reduction in output has a

favorable terms-of-trade effect. If the demand for the agriculture good is inelastic ($\eta < 1$) we reach the startling conclusion that the sacrifice becomes negative! This is, of course, the familiar proposition that aggregate farm income may be increased by reducing output. The *direct* gain in manufactured goods achieved by the rural sector through exporting an additional unit of labor is $\bar{W}_M N_M / N_u$, the expected urban wage. But additional migration, by increasing unemployment, reduces the earnings of *all* migrants already in the urban labor force by a factor $(1-R)$, where R is the fraction of the total urban labor force supplied by the rural sector.¹⁷

As long as $Pq'(1-\eta) < \bar{W}_M N_M / N_u (1-R)$ the welfare of the rural sector will be increased by allowing migration even though unemployment ensues and the economy as a whole sacrifices output. Since Pq' and $\bar{W}_M N_M / N_u$ are always positive and $R \leq 1$, additional migration will always benefit the rural sector when $\eta < 1$. In general, the lower is Pq' , η , or R and the higher is $\bar{W}_M N_M / N_u$, the more will the rural sector benefit from the opportunity to migrate.

From the foregoing, one can conclude that although migration restriction will improve aggregate welfare of the economy, given plausible values of η and R , substantial compensation to the rural sector will be required if it is not to be made worse off by removing the opportunity for free migration. The permanent urban labor force clearly will be made better off by becoming fully employed at the high

minimum wage while also being able to buy food at a lower price. Each unit of labor exported by the rural sector will similarly earn more but this gain will be offset by reduced total labor exports and lower agricultural prices. Whether or not this will be true depends, of course, on the values of the specific parameters of the economy. If η is sufficiently high, the rural sector could be made better off by restricting migration in the absence of compensation, but this seems very unlikely.

C. A Combination of Policies

It has been shown that either a limited wage-subsidy or a migration-restriction policy will lead to a welfare improvement. Which of the two policies will lead to the better position cannot be determined without knowing all the relevant parameters for a particular economy. It is clear, however, that neither policy alone is capable of moving the economy to the optimum that could be achieved with competitive wage determination (point L in Figure 2).

At first sight it may seem strange that with a single market failure, the wage level, a single policy instrument is unable to fully correct the situation.¹⁸ The reason is that the wage performs two functions in this model. It determines *both* the level of employment in the industrial sector *and* the allocation of labor between rural and urban areas. While a subsidy changes the effective wage for determination of industrial employment, so long as the wage actually received by workers exceeds agricultural earnings there will be migration and urban unemployment. Restriction of migration prevents the minimum wage having its effect on unemployment but does nothing to increase the level of industrial employment. Therefore, if the optimum position is to be achieved, a combination of both instruments will have to

¹⁷ If the urban unemployment were experienced only by migrants, this term would equal zero since the total amount of earnings through labor export would be constant. It can be positive only because the permanent urban labor force shares in unemployment, thereby reducing its share of the constant wage bill in the manufactured good industry. An interesting extension of the model would be to incorporate different employment probabilities for the permanent urban and migrant rural labor forces and then to check the sensitivity of results with our more simplified assumption of equal probabilities.

¹⁸ We wish to thank a referee of this *Review* for drawing this to our attention.

be used. In order to reach point L a wage subsidy must be instituted such that industrial employment will increase to the extent that with full employment the marginal product of labor will be equal in manufacturing and agriculture. The subsidy will be positive and equal to the difference between the minimum wage and marginal productivity. At that point $W_u^* = \bar{W}_M$ and $\bar{W}_M > Pq'$. Therefore, individuals would still find it in their interest to migrate and the point will not be attainable unless migration is restricted.

The agricultural sector has to be better off at L than at E since each additional unit of labor exported earns the full minimum wage, marginal productivity in agriculture is less than the minimum wage, and the price of the agricultural good rises. Whether the agricultural sector is better off at L than at D , however, depends again on the parametric values of the model.¹⁹ It can be stated with certainty that the amount of compensation needed to make the rural sector *no worse off* than at D will be less at L than at E , and, furthermore it should be easier to finance since total income is greater.

Even so the fiscal requirements of subsidy (or public enterprise losses) and compensation cannot be taken lightly.²⁰ A government may find it difficult to find

¹⁹ As drawn in Figure 2, L must represent a higher welfare level than D for the rural sector since P rises and the sector produces more of both goods. In fact if L lies along TS north of the ray going through D there will be an unambiguous sectoral welfare improvement. However, if L lies south of the ray on TS , the rural sector could be worse off than at D since P falls.

²⁰ This argument coincides with the statement by Stolper (p. 195), "It should be noted, however, that even at best the application of shadow prices leads to the substitution of one problem, the budget, for another one, an imperfect market."

We would not go as far as Stolper in rejecting out of hand any use of shadow pricing because of the fiscal implications. The general point is valid that one cannot disregard the consequences of implementation of shadow-price criteria if actual prices or wages continue to diverge from the shadow prices or wages.

nondistorting taxes capable of raising sufficient revenue. Perhaps a head-tax on all urban residents would be feasible although this too raises the question of how minimum wages are set (unions in tropical Africa have, in some cases, successfully fought to maintain the real after-tax wage). A tax on rural land is ruled out if there must be *net* compensation to the rural sector which, in the absence of pure profits in manufacturing, leaves an urban land tax as the remaining potential ideal tax.

All of the above suggests that altering the minimum wage may avoid the problems of taxation, administration, and interference with individual mobility attendant to the policy package just discussed. Income and wages policies designed to narrow the rural-urban wage gap have been suggested by D. P. Ghai, and Tanzania has formally adopted such a policy along with migration restriction. In the final analysis, however, the basic issue at stake is really one of political feasibility and it is not at all clear that an incomes policy is any more feasible than the alternatives.

APPENDIX I

Proof of Stability of Unemployment Equilibrium

In order to prove that our urban unemployment equilibrium is stable, we can differentiate ψ (equation (9)) with respect to N_u remembering that $dN_u = -dN_A$ according to (7). We therefore obtain

$$(1.1) \quad \frac{dN_u}{dN_u} = \psi'(\cdot) \left[-\frac{\bar{W}_M N_M}{(N_u)^2} + Pq'' + \frac{\partial P}{\partial X_A} (q')^2 \right].$$

Stability requires $d\dot{N}_u/dN_u < 0$ which is satisfied if

$$\frac{\partial P}{\partial X_A} < \frac{\frac{\bar{W}_M N_M}{(N_u)^2} - Pq''}{(q')^2}.$$

The right side of this inequality is unambiguously positive since $q'' < 0$. Hence our assumption that $\partial P / \partial X_A < 0$ will ensure stability and, indeed, is stronger than necessary. The adjustment mechanism may be made clear by the following phase diagram in which the function ψ is plotted. Its positive slope reflects the hypothesis that migration flows will increase with the magnitude of the urban-rural expected wage differential. In Figure 4, ψ is plotted under the assumption that $\psi(0) = 0$, hence the horizontal intercept is at the origin (in general the intercept would be α). Furthermore, we have arbitrarily assumed that ψ is a linear function. The arrows show the direction of adjustment in accordance with (1.1). If $\bar{W}_M N_M / N_u - P q' > 0$, then $\dot{N}_u > 0$ but we know that if $\dot{N}_u > 0$, the expected wage differential will decrease since $d\bar{N}_u / dN_u < 0$. Additional migration by increasing N_u without affecting N_M will reduce the expected urban real wage through increased unemployment. Concomitantly, the transfer of labor out of agriculture raises q' and reduced agricultural output also causes P to rise. Thus migration reduces the expected wage differential to zero and equilibrium is achieved when there is no further incentive for migration. See Todaro for a more detailed analysis of this process in a dynamic setting.

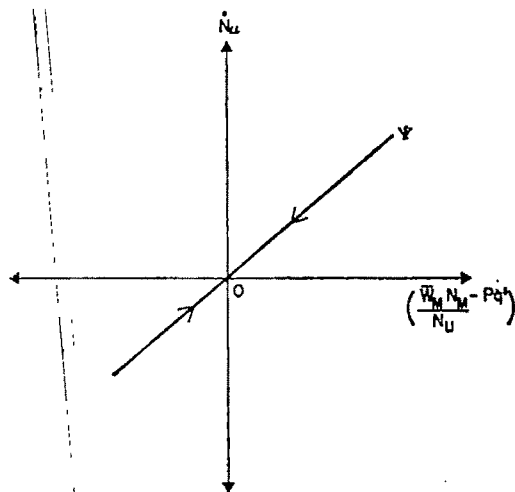


FIGURE 4.

APPENDIX II

Differentiating the equilibrium condition (8) with respect to N_M , recalling that $dN_u = -dN_A$, we obtain the expression

$$(II.1) \quad \frac{dN_u}{dN_M} = \frac{\frac{\bar{W}_M}{N_u} - q' \rho' \frac{f'}{X_A}}{\frac{\bar{W}_M N_M}{N_u^2} - \rho q'' + q' \rho' \frac{q' X_M}{X_A^2}}.$$

Defining the elasticity of demand for the agricultural good as

$$(II.2) \quad \eta_A = - \frac{\partial X_A}{\partial P} \cdot \frac{P}{X_A} = \frac{\rho X_A}{\rho' X_M},$$

(II.1) can be rewritten as

$$(II.3) \quad \frac{dN_u}{dN_M} = \frac{\frac{\bar{W}_M}{N_u} - \frac{\rho q' f'}{\eta_A X_M}}{\frac{\bar{W}_M N_M}{N_u^2} - \rho q'' + \frac{\rho (q')^2}{\eta_A X_A}}.$$

Differentiating the expression partially with respect to its various arguments it can be shown that dN_u / dN_M will vary directly with \bar{W}_M , N_M , η_A and inversely with ρ , q' , f' , N_u , and q'' . In general, the greater is the urban-rural wage differential, and the less sensitive are prices and marginal products in agriculture, the greater will be the migration induced by creation of an additional job. If the minimum wage exceeds agricultural earnings, (II.3) will generally be positive and, with parameter values relevant for many African economies, will exceed unity.

When $dN_u / dN_M > 1$, creation of an additional job at the minimum wage will increase the absolute level of unemployment although the *rate* of urban unemployment will have to fall. This can be seen by converting (II.3) to an elasticity measure.

$$(II.4) \quad \frac{dN_u}{dN_M} \cdot \frac{N_M}{N_u} =$$

$$\frac{\frac{\bar{W}_M N_M}{N_u^2} - \frac{N_M \rho q' f'}{N_u \eta_A X_M}}{\frac{\bar{W}_M N_M}{N_u^2} - \rho q'' + \frac{\rho (q')^2}{\eta_A X_A}} < 1$$

since $q'' < 0$.²¹ To give an example of what this means, suppose that an economy initially has an urban unemployment rate of 25 percent. If in response to the creation of 100 additional industrial jobs, 125 additional individuals migrate to the urban area, the absolute number unemployed increases by 25 although the unemployment rate will drop, since the marginal unemployment rate is only 20 percent.

APPENDIX III

If minimum wages are maintained and migration takes place in accordance with equation (8), aggregate welfare will be maximized if the following Lagrangean expression is maximized:

$$\begin{aligned} \Omega = & U(X_A, X_M) \\ & + \lambda_1 [q(\bar{N} - N_u) - X_A] \\ & + \lambda_2 [f(N_M) - X_M] \\ (III.1) \quad & + \lambda_3 \left\{ \rho \left(\frac{f(N_M)}{q(\bar{N} - N_u)} \right) \right. \\ & \left. \cdot q'(N - N_u) - \frac{\bar{W}_M N_M}{N_u} \right\} \end{aligned}$$

where U is the social welfare function and the succeeding terms are the constraints imposed by equations (1), (2), and (8) (recall that $N_A = \bar{N} - N_u$ from equation (7)).

Maximizing (III.1) we get the following first-order conditions:

$$(III.2) \quad \frac{\partial \Omega}{\partial X_A} = \frac{\partial U}{\partial X_A} - \lambda_1 = 0$$

$$(III.3) \quad \frac{\partial \Omega}{\partial X_M} = \frac{\partial U}{\partial X_M} - \lambda_2 = 0$$

$$(III.4) \quad \frac{\partial \Omega}{\partial N_u} = -\lambda_1 q' + \lambda_3 \left[\rho' \frac{f q'}{q^2} - \rho q'' + \frac{\bar{W}_M N_M}{N_u^2} \right] = 0$$

$$(III.5) \quad \frac{\partial \Omega}{\partial N_M} = \lambda_2 f' + \lambda_3 \left[\rho' \frac{f' q'}{q} - \frac{\bar{W}_M}{N_u} \right] = 0$$

and the $\partial \Omega / \partial \lambda_i = 0$ ($i = 1, 2, 3$) which ensures that the constraints hold.

Substituting (III.2) and (III.3) into (III.4) and (III.5) we get

$$(III.6) \quad \frac{\frac{\partial U}{\partial X_M} f'}{\frac{\partial U}{\partial X_A} q'} = \frac{\frac{\bar{W}_M}{N_u} - q' \rho' \frac{f'}{q}}{\frac{\bar{W}_M N_M}{N_u^2} - \rho q'' + q' \rho' \frac{f q'}{q^2}}$$

We know that in equilibrium $(\partial U / \partial X_M) / (\partial U / \partial X_A) = 1/P$ and it has been shown in Appendix II that the right-hand side of (III.6) is equal to dN_u / dN_M . Therefore (III.6) can be rewritten as

$$(III.7) \quad f' = P q' \frac{dN_u}{dN_M},$$

which is the condition used in the text to determine the optimal wage subsidy.

Condition (III.7) can also be written as

$$(III.8) \quad -P = \frac{-f'}{q' \frac{dN_u}{dN_M}} = \frac{dX_M}{dX_A}.$$

We know that $-P$ is equal to the marginal rate of substitution between the two commodities and dX_M / dX_A is the marginal rate of transformation. Hence (III.8) states the familiar condition for optimality: equate marginal rates of substitution and transformation. dX_M / dX_A is the slope of the line DK in Figure 2 and it clearly will be nega-

²¹ We are grateful to Peter Diamond for deriving this expression.

tive. However, its derivative with respect to N_M ,

$$(III.9) \quad \frac{d\left(\frac{dX_M}{dX_A}\right)}{dN_M} = \frac{-q' \frac{dN_u}{dN_M} f'' - f' \left(\frac{dN_u}{dN_M}\right)^2 q'' + f' q' \frac{d^2 N_u}{dN_M^2}}{\left(q' \frac{dN_u}{dN_M}\right)^2}$$

is of indeterminate sign since f'' , $q'' < 0$ and $d^2 N_u / dN_M^2$ will generally be negative as well. (III.9) must be positive if the effective production possibilities frontier (DK) is to be convex, a condition that is likely to hold but the possibility of concavity as full employment is approached must be considered. The slope of DK in Figure 2 seems plausible on a priori grounds.

REFERENCES

- P. K. Bardham, "Factor Market Disequilibrium and the Theory of Protection," *Oxford Econ. Pap.* (New Series), Oct. 1964, 16, 375-88.
- E. J. Berg, "Wage Structure in Less Developed Countries," in A. D. Smith, ed., *Wage Policy Issues in Economic Development*, London 1969.
- A. Callaway, "From Traditional Crafts to Modern Industries," *ODU: University of Ife Journal of African Studies*, July 1965, 2.
- S. Chakravarty, "The Use of Shadow Prices in Programme Evaluation," in Rosenstein-Rodan, ed., *Capital Formation and Economic Development*, London 1964.
- Y. S. Cho, *Disguised Unemployment in Developing Areas, with Special Reference to South Korean Agriculture*, Berkeley 1960.
- R. S. Eckaus, "The Factor-Proportions Problem in Underdeveloped Areas," *Amer. Econ. Rev.*, Sept. 1955, 45, 539-65.
- J. Erickson, "Wage Employment Relationships in Latin American Industry: A Pilot Study of Argentina, Brazil, and Mexico," International Labour Office, 1969, type-script.
- J. Fei and G. Ranis, *Development of the Labor Surplus Economy*, Illinois 1964.
- D. P. Ghai, "Incomes Policy in Kenya: Need, Criteria and Machinery," *East Afr. Econ. Rev.*, June 1968, 4, 19-35.
- J. Gugler, "The Impact of Labour Migration on Society and Economy in Sub-Saharan Africa. Empirical Findings and Theoretical Considerations," *African Social Research*, Dec. 1968, 6, 463-86.
- E. E. Hagen, "An Economic Justification of Protectionism," *Quart. J. Econ.*, Nov. 1958, 72, 496-514.
- J. R. Harris and M. P. Todaro, "Urban Unemployment in East Africa: An Economic Analysis of Policy Alternatives," *East Afr. Econ. Rev.*, Dec. 1968, 4, 17-36.
- and —, "Wages, Industrial Employment, and Labour Productivity: The Kenyan Experience," *East Afr. Econ. Rev.* (New Series), June 1969, 1, 29-46.
- J. P. Henderson, "Wage Policy in Africa," Paper prepared for delivery at the African Conference on Economics, Temple University, mimeo, April 1968.
- C. R. Hutton, "The Causes of Labour Migration," in Gugler, ed., *Urbanization in Sub-Saharan Africa*, Kampala 1969.
- C. H. C. Kao, K. R. Ansel, and C. K. Eicher, "Disguised Unemployment in Agriculture: A Survey," in C. K. Eicher and L. W. Witt, eds., *Agriculture in Economic Development*, New York 1964, 129-44.
- J. M. Katz, "Verdoorn Effects; Returns to Scale, and the Elasticity of Factor Substitution," *Oxford Econ. Pap.*, Nov. 1968, 20, 342-52.
- L. Lefebvre, "Planning in a Surplus Labor Economy," *Amer. Econ. Rev.*, June 1968, 58, 343-73.
- W. A. Lewis, "Economic Development with Unlimited Supplies of Labour," *Manchester Sch. Econ. Soc. Stud.*, May 1954, 22, 139-91.
- I. M. D. Little, "The Real Cost of Labour, and the Choice Between Consumption and

- Investment," in P. N. Rosenstein-Rodan, ed., *Pricing and Fiscal Policies: A Study in Method*, Cambridge 1964, 77-91.
- and J. A. Mirrlees, *Manual of Industrial Project Analysis*, Vol. II, "Social Cost Benefit Analysis," Paris 1969.
- B. F. Massell and J. Heyer, "Household Expenditure in Nairobi: A Statistical Analysis of Consumer Behaviour," *Econ. Develop. Cult. Change*, Jan. 1969, 17, 212-34.
- L. G. Reynolds, "Wages and Employment in a Labor-Surplus Economy," *Amer. Econ. Rev.*, Mar. 1965, 55, 19-39.
- W. F. Stolper, *Planning Without Facts: Lesson's in Resource Allocation from Nigeria's Development*. Cambridge 1966.
- M. P. Todaro, "A Model of Labor Migration and Urban Unemployment in Less Developed Countries," *Amer. Econ. Rev.*, Mar. 1969, 59, 138-48.
- Nigeria, *Report of the Commission on the Review of Wages, Salary and Conditions of Service of the Junior Employees of the Governments of the Federation and in Private Establishments 1963-64*.

On the Economic Welfare of Victims of Automobile Accidents

By R. A. HOLMES*

The economic welfare of victims of automobile accidents depends in part on the amount of compensation received, and in part on the time lapse between accident and compensation. Fulfillment of these two dimensions of economic welfare varies widely, and in large part, nonrandomly among accidents. Severe hardships may result when the amount of compensation falls far short of the economic cost of an accident or when the receipt of compensation is long delayed. The purpose of this study is to assess the present system of compensating victims of traffic accidents by explaining and interpreting the non-random component in both the amounts of compensation received relative to losses suffered, and the time lag from accident to compensation.¹

I. *The Survey Design*²

The universe for this sample study includes all residents of British Columbia, Canada, who were involved in traffic ac-

cidents in the calendar year 1963.³ Simple random samples were obtained from accident reports⁴ and the survey was designed to yield 2.5 percent of all noninjury cases, 10 percent of all nonfatal injuries, and a census of fatality cases occurring in the 13 cities, 12 municipalities, 5 villages, and 5 R.C.M.P. detachments included in the study.⁵

In fatality cases, relatives of the victims were personally interviewed. In all other cases, the first stage consisted of mail questionnaires seeking information on property damage and compensation; medical, hospital, and other expense and compensation; income loss and compensation; legal proceedings; attitudes; and other related matters. The serious injury cases among the respondents to the mail questionnaire were then defined as those involving one or more of the following three characteristics:

1. Medical expenses of \$500 or more
2. Loss of three or more weeks of work
3. Permanent physical impairment affecting ability to work.

All other respondents to the mail questionnaire were defined as minor injury cases.⁶

*The author is professor of economics at Simon Fraser University.

¹ The author is indebted to the British Columbia Royal Commission on Automobile Insurance for financial assistance provided, to Dr. Sylvia Ostry for unpublished data on work life expectancy of Canadian males, to Dr. R. G. Wilson for time spent in gaining the support of the B.C. Medical Association, to nine able and conscientious research assistants (Miss M. I. Edgar, Miss E. E. Mackenzie, Miss G. E. Mathias, Miss M. K. Swain, R. D. Carlson, K. K. Foo, H. H. Jensen, H. K. Larsen, and H. B. Yee), and most especially, to the respondents for their cooperation. Helpful comments have also been made by James Morgan and an anonymous referee.

² The procedural guidelines for this study were laid down in a pioneering work in this field, see A. F. Conard, et al. A flow diagram providing a detailed description of our survey design is available on request.

³ The year 1963 may seem to be unnecessarily far in the past, but a more recent year would not have allowed sufficient time for the settlement of cases involving litigation.

⁴ The accident reports were completed by attending policemen and filed with the Motor Vehicles Branch in Victoria, B.C.

⁵ Searches through telephone and city directories, as well as files of persons currently holding British Columbia drivers licenses were undertaken to identify current addresses in cases of address change since 1963.

⁶ With minor injury cases, the questionnaires returned sometimes provided all of the information needed, but frequently the questionnaire was returned

Personal interviews were conducted with serious injury cases as well as with relatives of persons killed. These interviews were used to complete any questions that had not been answered adequately on the mail questionnaire, to collect more detailed information on medical expenses, income losses, and attitudes, to obtain the names and addresses of physicians, hospitals, and lawyers involved, and to obtain authorization for the release of pertinent information held by other persons or institutions, (particularly hospitals, physicians, and lawyers).

II. *The Response*

Our total sample size was 2,765 persons, of whom 45 percent provided complete information. Only 29 percent did not return our mail questionnaires or refused interviews, and the remainder consist of 10 percent who attempted but were unable to provide complete information, and 15 percent whose addresses (or the addresses of survivors) were unknown, or who had died since, from causes other than the accident. Our nonresponse rates (including those providing incomplete information) are 54 percent of persons involved in nonfatal accidents, 55 percent of the fatality cases, and 67 percent of persons suffering serious or minor injuries in accidents involving fatalities. Although the latter group may be differently and more seriously affected by nonresponse bias, it is a very small part of the population and of the sample included in this study. Nonresponse rates for serious and minor injuries in nonfatal accidents and for accidents involving fatali-

ties, which together constitute 95 percent of the cases considered, are almost identical, and these groups appear not to be differently affected by nonresponse bias. Moreover, the nonresponse rates are low compared to those usually encountered when data are collected by interview; this undoubtedly reflects publicity given to the enquiry and widespread public interest.

III. *Measurement: Problems and Procedures*

In order to measure the severity of accidents and the adequacy of compensation received, it is necessary to devise measures of economic loss and compensation. This is particularly difficult in the case of losses. With compensation,⁷ the difficult conceptual problem concerns the handling of legal costs which are incurred in obtaining compensation. These legal costs could be included as part of total costs in which case they would not be deducted from total compensation, or they could be excluded from costs and deducted from total compensation. The latter course is followed in this study on the grounds that money used for collection expenses is not compensation as far as the victim is concerned. Consequently, our estimates of total costs exclude legal costs.

Conceptual problems arise more frequently in the measurement of economic losses. Many of these problems defy perfect solution and our estimates are based on assumptions which are consciously designed to yield underestimates of economic losses. In other words, our methods conform to established statistical procedure by giving the benefit of the doubt to the compensation system which is under test.

with some questions left unanswered or with inconsistent answers. In these cases, the respondent was phoned and complete information obtained, if possible. Of course, there were cases, (fatality and serious injury as well as minor injury cases), where complete information could not be obtained even though the respondent was willing to cooperate, and in these cases the questionnaires were classified as incomplete and were not included among those subsequently coded and analyzed.

⁷ Compensation includes amounts received from the victims' own insurance (automobile, life, medical, hospital, and accident), as well as the amounts received from the other person or persons responsible or their insurance companies, or from such sources as the Workmen's Compensation Board, The Traffic Victims Indemnity Fund, employers (sick leave), or welfare payments.

Economic loss is defined to consist of expenses incurred, expected future expenses, and income loss.⁸ The latter is most difficult to measure.⁹ Income loss is the largest loss of persons who are seriously injured or killed in automobile accidents, and it is difficult to measure accurately because it requires forecasts of future occupation, education level, employment status (i.e. whether or not employed), rate of pay, maintenance costs, and working life, as well as the choice of an appropriate discount rate.

The occupation and education levels of employed persons who are injured are taken as found, and it is assumed that these would not have changed in the absence of the accident. Annual salaries at the time of the accident are used as "bench marks" which are projected according to two criteria:

1. Movements through a stable salary structure. We estimate how, under stable conditions, income changes as experience increases, from census data which shows for given age, education, and sex, how income changes as age changes.

2. Shifts in the salary structure. We assume that salary structures shift upward by rates equal to long-term productivity growth of the economy, which we conservatively estimate as 2 percent annually. This means that our estimates of future income losses are estimates of real income which incorporate no inflationary effects.

With students, both the level of and future changes in earnings are projected from census data¹⁰ on the assumption of no

future change in education level. With housewives, income losses are derived from conservative estimates of replacement cost.¹¹

The employment status of all individuals is projected as if, in the absence of the accident, there would have been no change from the time of the accident. Those who are unemployed are estimated to suffer no future income loss because of the accident, while those who are employed are estimated to suffer no future unemployment. If a person is found to be seasonally unemployed at the time of the accident, we estimate that, in the absence of the accident, he would have experienced as much seasonal unemployment per year as in the year previous to the accident. Although our estimates of future income loss for individual cases suffer from the impossibility of predicting future employment status for individuals, our estimates of aggregate future income loss, in which we are primarily interested, are simply based on the assumption that future unemployment rates in British Columbia remain at their 1963 level (6.4 percent).

Estimates of future income loss also depend on whether or not maintenance or subsistence costs are deducted in the case of fatalities. The question here is whether all consumption should be treated as income which adds to total satisfaction, or whether part of consumption should be treated as subsistence cost or cost of production. If all consumption is treated as

⁸ Expenses incurred include medical and hospital expense, repair of automobiles, and other expenses resulting from the accident. Expected future expenses consist only of medical expenses in this study, and these are included only in cases where they appear certain to be incurred.

⁹ A flow diagram providing a detailed description of our methods of measuring income loss is available on request.

¹⁰ A work life expectancy equal to that for 15 year olds is used for children under 15 years of age.

¹¹ In cases of housewives with children under 12 years of age, replacement cost is estimated to involve the services of a live-in housekeeper as well as a day worker on the seventh day of the week as long as the housewife's youngest child remains under 12 years of age. With a monthly salary of \$250 for the housekeeper and a daily salary of \$10 for the day worker, the annual replacement cost is \$3520 which we round downward to \$3500. In cases of disabling injuries to housewives with no children under 12 years of age, we assume replacement cost to involve a housekeeper for 4 hours per day at \$1.25 per hour which gives an annual replacement cost of \$1725, and again we round down to \$1700.

income, then no distinction is made between the income loss resulting when an individual is killed in an accident, and when he is permanently and completely disabled, even though in the latter case the family or society would continue to provide subsistence costs. One could reasonably argue that if the victim remains alive, this is a source of satisfaction to him, his family, and to society, and that one should make no deduction for subsistence costs. However, the question is contentious and the desire to underestimate rather than overestimate future income loss determines our procedure. Only earnings in excess of \$75 per month are considered income, and conversely, the first \$75 earned each month are treated as subsistence costs which do not add to total satisfaction.¹²

Work life expectancies and discount rates are also required in estimating future income losses. We have used estimates of expected length of work life for Canadian males and U.S. females.¹³ The discount rate used is 7.5 percent which corresponds to about 9 percent on real markets with price inflation of 1.5 percent annually.¹⁴ This is a much higher rate of discount than is ordinarily used, being an expected stock rather than bond yield. However, the higher discount rate is consistent with our

intention to underestimate economic loss in order to ensure that any biases in our estimates work to the advantage of the present compensation system.

Subsequently, comparisons will be made of our estimates of economic loss and compensation in order to assess the adequacy of the present system of compensating victims of traffic accidents. The results of this assessment are strongly biased in favour of the present compensation system because, as previously shown, our estimates of losses are biased downwards. The sources of this bias, which should be kept clearly in mind when interpreting the results, are the following:

1. No allowances are made for pain and suffering in calculating economic loss.

2. The income losses of all persons (children and adults) are estimated on the assumption that their education level would not, in the absence of the accident, have advanced beyond the level attained at the time of the accident.

3. The overall unemployment rate implicit in our estimates of future income losses is the British Columbia rate for 1963 (6.4 percent) which is abnormally high.

4. Although there is reason to doubt whether any allowance should be made for subsistence costs in estimating future income loss, we make an allowance of \$75 per month.

5. Future growth in average incomes is projected at an annual rate of 2 percent per year which is a conservative estimate of overall increases in productivity.

6. Future income losses are discounted at a 9 percent annual rate (assuming inflationary effects of 1.5 percent) which is an expected equity rather than debt yield.

IV. *The Model and Two-Stage Least Squares Estimates*

The a priori restrictions imposed in this study are that compensation received by victims of traffic accidents is a linear

¹² This \$75 figure is equal to the old age pension provided by the Federal government in 1963, and seems therefore to be a generous allowance for subsistence costs, particularly when one bears in mind that there is considerable doubt whether any such deduction should be made.

¹³ The Canadian male working life estimates used were prepared by the Department of Labour, and examples of them are published in *The Labour Gazette*, March 1966 (p. 99). Unfortunately, no similar tables exist for Canadian females, and we use U.S. data on the working lives of females from BLS. We estimate the working life of housewives from standard life tables for Canadian women.

¹⁴ Since the effects of inflation are not incorporated in our projections of future earnings, consistency requires that we remove inflationary effects from our discount rate.

stochastic function of variables representing the stage of litigation, the type of injury and accident, and the amount of economic loss, while the delay to compensation is a linear stochastic function of variables representing the amount of compensation received, the stage of litigation, and the type of injury and accident. More precisely, our model stipulates:

$$(1) \quad \begin{aligned} y_{1r} = & \gamma_{11} + \gamma_{12}x_{2r} + \gamma_{13}x_{3r} + \gamma_{14}x_{4r} \\ & + \gamma_{15}x_{5r} + \gamma_{16}x_{6r} + \gamma_{111}x_{11r} \\ & + \gamma_{112}x_{12r} + \gamma_{113}x_{13r}^2 + u_{1r} \end{aligned}$$

$$(2) \quad \begin{aligned} y_{2r} = & \gamma_{21} + \beta_{21}y_{1r} + \beta_{22}y_{1r}^2 + \gamma_{22}x_{2r} \\ & + \gamma_{23}x_{3r} + \gamma_{24}x_{4r} + \gamma_{25}x_{5r} \\ & + \gamma_{26}x_{6r} + \gamma_{27}x_{7r} + \gamma_{28}x_{8r} \\ & + \gamma_{29}x_{9r} + \gamma_{210}x_{10r} + u_{2r} \end{aligned}$$

where:

$r = 1, \dots, N$

N = population size

y_{1r} = compensation (in dollars) received by the r th individual

y_{2r} = delay (in months) to final compensation for the r th individual

$x_{2r} = 1$ if the r th individual did not hire a lawyer
= 0 otherwise

$x_{3r} = 1$ if the r th individual hired a lawyer but settled without either suit or trial
= 0 otherwise

$x_{4r} = 1$ if the r th individual hired a lawyer and filed suit but settled without a trial
= 0 otherwise

$x_{6r} = 1$ if the r th individual hired a lawyer, filed suit and went to trial
= 0 otherwise

$x_{9r} = 1$ if the r th individual sustained minor injuries in an accident involving no fatalities
= 0 otherwise

$x_{7r} = 1$ if the r th individual sustained

serious injuries in an accident involving no fatalities

= 0 otherwise

$x_{8r} = 1$ if the r th individual is a fatality case

= 0 otherwise

$x_{9r} = 1$ if the r th individual sustained serious injuries in an accident involving some fatalities

= 0 otherwise

$x_{10r} = 1$ if the r th individual sustained minor injuries in an accident involving some fatalities

= 0 otherwise

$x_{11r} = 1$ if the r th individual is not a fatality case

= 0 otherwise

x_{12r} = amount of economic loss (in dollars) sustained by the r th individual

u_{1r}, u_{2r} = disturbance terms

Equation (1) specifies that the compensation received by the r th individual depends on the stage of litigation as represented by variables x_2 , x_3 , x_4 , and x_6 ; whether or not the r th individual was fatally injured (other injury classes were tested but found to be insignificant); and on economic loss (whose partial relation with compensation received is a second degree quadratic). Equation (2) specifies that the delay to final compensation for the r th individual depends on the stage of litigation; whether the injury is minor, serious, or fatal; whether or not the accident involved fatalities; and the amount of compensation received (whose partial relation with y_2 is also specified to be a second degree quadratic).¹⁵

The existence of the endogenous variable y_1 on the right-hand side of equation

¹⁵ We include the y_{1r}^2 term in equation (2) because of the belief that the delay to final compensation at first increases rapidly with the amount of compensation received, but reaches its maximum at relatively low amounts of compensation.

(2) leads to bias in the ordinary least squares estimates of the parameters of that equation. To obtain consistent estimates we employ two-stage least squares, eliminating x_2 and x_{11} from (1) and x_2 and x_8 from (2). This yields the following estimates, where t -ratios are in parentheses:

$$\begin{aligned} (3) \quad \hat{y}_{1r} = & 19.26 + 630.4x_{3r} + 1167x_{4r} \\ & (2.15) \quad (3.04) \\ & + 235.3x_{5r} - 1751x_{8r} \\ & (0.31) \quad (-3.63) \\ & + 0.4871x_{12r} - 2.997(10^{-6})x_{12r}^2 \\ & (19.67) \quad (-14.00) \\ R = & 0.63 \quad F = 135.5 \end{aligned}$$

$$\begin{aligned} (4) \quad \hat{y}_{2r} = & 2.46 + 0.00096\hat{y}_{1r} \\ & (4.33) \\ & - 0.040(10^{-6})\hat{y}_{1r}^2 + 1.35x_{3r} \\ & (-3.05) \quad (2.77) \\ & + 13.50x_{4r} + 10.15x_{5r} \\ & (19.85) \quad (8.41) \\ & + 2.91x_{7r} + 1.25x_{8r} \\ & (4.53) \quad (1.74) \\ & + 2.57x_{9r} - 1.16x_{10r} \\ & (1.78) \quad (-1.68) \\ R = & 0.67 \quad F = 115.1 \end{aligned}$$

where:

$$r = 1, \dots, 1253$$

$$\hat{y}_{jr} = y_{jr} - e_{jr}$$

e_{jr} = least squares residuals from the j th equation

Equation (1) was also fitted separately for the minor injury cases to obtain:

$$\begin{aligned} (5) \quad \hat{y}_{1r} = & -10.97 + 45.12x_{3r} + 461.1x_{4r} \\ & (1.68) \quad (10.06) \\ & + 288.2x_{5r} + 0.8132x_{12r} \\ & (2.80) \quad (41.41) \\ R = & 0.81 \quad F = 506.8. \end{aligned}$$

The variable x_{12r}^2 was dropped from (5) because its coefficient was found to be insignificantly different from zero in this case. The other exogenous variables (x_{7r}

$\dots x_{10r}$) are irrelevant when considering the minor injury cases separately.

For ease of interpretation, the omitted variables are returned to equations (3), (4), and (5) by a conversion which sets the weighted average of the coefficients of each set of dummy variables (rather than the coefficient of the excluded variable) equal to zero.¹⁶ These conversions give the following estimates:

$$\begin{aligned} (6) \quad \hat{y}_{1r} = & 83.36 - 149.9x_{2r} + 480.5x_{3r} \\ & + 1017.1x_{4r} + 85.4x_{5r} \\ & - 1665.2x_{8r} + 85.8x_{11r} \\ & + 0.4871x_{12r} - 2.997(10^{-6})x_{12r}^2 \end{aligned}$$

$$\begin{aligned} (7) \quad \hat{y}_{2r} = & 3.68 + 0.00096\hat{y}_{1r} \\ & - 0.040(10^{-6})\hat{y}_{1r}^2 - 1.0x_{2r} \\ & + 0.4x_{3r} + 12.5x_{4r} + 9.2x_{5r} \\ & - 0.2x_{8r} + 2.7x_{7r} + 1.0x_{8r} \\ & + 2.4x_{9r} - 1.4x_{10r} \end{aligned}$$

$$\begin{aligned} (8) \quad \hat{y}_{4r} = & 22.18 - 33.2x_{2r} + 12.0x_{3r} \\ & + 428.0x_{4r} + 255.1x_{5r} \\ & + 0.8132x_{12r} \end{aligned}$$

V. Conclusions

We consider in this section five conclusions that we draw from the regression estimates previously described and from a few supplementary descriptive statistics which are described below.

1. *The overall amount of compensation provided to the victims of automobile accidents is inadequate.*

Economic losses averaged \$2200 as compared to average compensation of only \$700. Minor injury cases with average compensation equal to 80 percent of average economic losses and an average net cost to the individual of only \$50 may

¹⁶ See J. N. Morgan for the method of calculating these coefficients.

be adequately compensated, but serious injury cases and the families of fatalities whose average compensations are only 45 percent and 20 percent, respectively, of average economic loss, and whose average net costs are \$4000 and \$23,000, respectively, are not. This conclusion is a judgment based not only on these average figures but also on the observation of individual cases of extreme hardship resulting from traffic accidents.¹⁷

2. The compensation provided is unfairly distributed by amount and type of economic loss.

Equation (8) shows that in minor injury cases, expected compensation is about 80 percent of economic loss for all values of economic loss. Of course, the range of economic loss is relatively small (0 to \$3000) in minor injury cases. In serious injury cases where much larger amounts of economic loss are encountered, the bias in the distribution of compensation against the large loss cases is shown by the highly significant negative coefficient of x_{12} in equations (3) and (6). These equations reveal that in serious injury cases where no lawyer is consulted (30 of 90 cases), the recovery rate (i.e. compensation as a percentage of economic loss) declines steadily from about 50 percent for small amounts of economic loss to only 26 percent for economic loss of \$75,000. In other words, the individual's expected net cost is only \$1300 in a \$2500 accident but it is \$55,000 in a \$75,000 accident.¹⁸

This discrimination by amount of loss is closely related to discrimination by kind of

loss. Since 97 percent of the losses of fatality cases, and two-thirds of the losses of serious injury cases are income losses, whereas 85 percent of the losses of minor injury cases are automobile damage, the differences between the recovery rates for the three groups (80, 45, and 20 percent respectively) suggests very clearly that the present compensation system discriminates between automobile damage on the one hand, and personal injuries resulting in income loss on the other hand.¹⁹

The reason for this bias by kind of loss lies primarily in the difference between two-party insurance which covers automobile damage and three-party insurance for income loss. Accordingly, an insured is protected by his own insurance company against the amount of damage to his automobile in excess of the deductible provision, without regard to fault. With income loss resulting from personal injuries, however, three-party insurance prevails, and one's own automobile insurance company provides protection only against one's liability for the losses of others. If compensation for income loss is to be obtained from the automobile insurance industry, it must be obtained from the other person's insurance company, and that requires proof that the other person is at fault, and is therefore at least partly

¹⁷ For example, in one of the fatality cases involving death of a breadwinner, economic losses totalled \$155,000 while total compensation to survivors was only \$250. It is also interesting to note that a discount rate greater than 30 percent would need to be employed before full compensation for economic loss would be shown.

¹⁸ This result is obtained from equation (6) by solving for y_{1r} when $x_{2r} = x_{11r} = 1$, $x_{3r} = x_{4r} = x_{6r} = x_{8r} = 0$, and $x_{12r} = \$2,500$ and \$75,000 respectively.

¹⁹ This bias of the system against large income loss cases exists as well in cases where tort settlement is received. For example, in nonfatal accident serious injury cases involving tort settlement, the median values of net tort settlement as a percent of economic loss decline steadily from 107 percent to 13 percent as the amount of economic loss increases from under \$1000 to \$10,000 or more. Similarly in fatality cases, these medians are 26 percent for losses under \$3000 but only 11 percent for losses in the \$3000 to \$49,999 range. It is interesting to note that the median values of gross tort settlement as a percentage of expenses incurred remains fairly constant over all expense classes for both serious and minor injuries in nonfatal accidents (about 160 percent and 90 percent respectively) and for fatality cases (about 100 percent), which implies that persons who receive tort settlements receive amounts proportional to their expenses, but not to income loss.

responsible for the income losses flowing from the accident.

3. *The distribution of compensation discriminates against the survivors of fatality cases.*

Equation (3) shows that after taking account of both the stage of litigation and the amount of economic loss, the expected compensation to relatives of persons killed in accidents is about \$1750 below that for minor or serious injury cases. The additional expected return to fatality cases is shown to be -\$1665 in equation (6), while that for nonfatality cases is +\$86. This difference is largely due to the greater relative importance in fatality cases of income loss which has lower probability of recovery than either medical expense or the expense of automobile damage.²⁰

4. *Lawyers probably contribute significantly to the net tort settlements received by their clients and therefore play an important role in the compensation system.*

Equation (6) shows additional expected returns of -\$150 in cases involving no lawyer, +\$480 in cases involving lawyer but no suit, +\$1017 in cases involving suit but no trial, and +\$85 in cases involving trial. It is important to bear in mind that in this study legal expenses are a deduction in the measurement of compensation so that these measured contributions of lawyers are contributions over and above the cost of their services. Even in the minor injury cases, equation (8) shows positive expected contributions from lawyers of \$12, \$428, and \$255, respectively, for the three successive stages of litigation involving lawyers, whereas the effect of no lawyer is negative (-\$33). The expected

return from the lawyer is then positive for every group considered and significantly different from cases involving no lawyer in 5 of the 6 tests conducted in equations (3) and (5). It is only in cases involving trial (equation (3)) where the expected difference drops to +\$235 that the difference from cases involving no lawyer is statistically insignificant at the .05 level in a one-tail test. All other tests show a significant increase in average return when a lawyer is employed. Equations (6) and (8) indicate that lawyers contribute most to their clients in cases involving suit but no trial, and that going to trial produces, on average, a negative marginal net benefit. Overall, the expected return with suit is \$1167 higher than for cases with no lawyer, but cases proceeding on to trial have an expected return \$932 below that peak. With minor injuries, cases with suit but no trial have an expected return \$461 higher than those with no lawyer, but the expected return to cases involving trial is \$173 below those involving suit but no trial, although it remains \$288 above those involving no lawyer.

These results suggest that too many cases are taken to trial, but they also establish the need of persons involved in traffic accidents, or their survivors, for lawyers in obtaining compensation. Unfortunately, this need is not usually recognized. The *beta* coefficient²¹ for the stage of litigation variables in equation (6) is only 0.085 as compared to 0.43 for the type of accident and injury dummies. The relatively low proportion of variance in compensation accounted for by the type of litigation dummies reflects the large number of persons who do not hire lawyers, probably because of fear of legal costs and underestimates of the effectiveness of lawyers in obtaining compensation for

²⁰ The difference may also be partly due to the fact that in some fatality cases (children for example) there are no identifiable dependents, and therefore no one to compensate for economic losses.

²¹ See J. N. Morgan for the method of calculating these *beta* coefficients.

economic losses resulting from traffic accidents.²²

5. Long delays from the time of accident to final compensation are commonplace. These delays frequently create serious financial problems for the person involved or his family, and this happens most frequently in cases involving suit and in cases where the amount of settlement is large. Although our previous results suggest that suit may be necessary to obtain adequate compensation, this avenue is open only to those who can wait an average of 20 to 30 months for compensation. Consequently, the amount and adequacy of the compensation received may frequently depend as much on ability to wait as on the validity of the case.

Equation (5) shows that the delay to final compensation depends significantly on the amount of compensation received, the stage of litigation involved, and the type of injury sustained. Least important, with a *beta* coefficient of 0.13, are the type of injury dummies. Equation (7) indicates negative additional expected delays for minor injuries in both nonfatal and fatal accidents (-0.2 and -1.4 months, respectively),²³ but in fatality and the two serious injury cases, they are positive (1.0, 2.7, and 2.4 months, respectively).

The litigation dummies have a very important effect on the delay to final compensation; their *beta* coefficient is 0.51

²² Only 47 percent of the serious injury cases, 23 percent of the survivors of fatality cases, 21 percent of minor injuries in fatal accidents, and 7 percent of minor injuries in nonfatal accidents were found to have an opinion when asked whether insurance companies will offer a larger settlement if the claimant is represented by a lawyer. Of those who do have opinions, 81 percent think that larger settlements are obtained with the assistance of lawyers, but this leaves 19 percent with the opposite opinion together with the majorities of each injury group having no opinion at all.

²³ The faster compensation for minor injuries in nonfatal accidents may reflect the desire of insurance companies to settle quickly and avoid litigation in the more serious accidents.

indicating that the stage of litigation accounts for about four times as much of the variance in delay to final compensation as the type of injury. Whereas the additional expected delay is -1.0 month for cases involving no lawyer and $+0.4$ months for cases involving a lawyer but no suit, it rises to 12.5 months with suit and to 9.2 months with trial. Obviously, if one is to obtain the higher return previously shown to be associated with suit, one must be able to wait substantially longer for final payment.

The additional delay with suit as compared to cases involving no lawyer will be longer than the 13.5 months differential in their regression coefficients in equation (7), because the time lag between accident and final compensation also depends on the amount of compensation received. This partial relationship (a second-degree quadratic) is highly significant (equation (4)), and the *beta* coefficients for β_1 and β_1^2 are relatively large (0.37 and 0.24 respectively). The nature of the relationship indicates that the resistance of the insurance industry rises very rapidly with compensation, but that most of the cases considered were settled in 30 months or less.

When we put together the various parameter estimates shown in equation (7), we see that the cases which suffer the most delay are those involving serious injury in nonfatal accidents, suit, and a relatively large amount of compensation. The expected delay for these cases range from about 18 months to more than 2 years as the amount of compensation increases over the \$1000 to \$12,000 range. We conclude therefore that some persons are probably forced to settle for inadequate compensation simply because they are unable to endure the long wait involved in obtaining more adequate reparation through the courts.

REFERENCES

- A. F. Conard, J. N. Morgan, R. W. Pratt, Jr., C. E. Voltz, and R. M. Bombaugh, *Automobile Accident Costs and Payments: Studies in the Economics of Injury Reparation*, Ann Arbor 1964.
- J. N. Morgan, *A Note on the Interpretation of Multiple Regression Using Dummy Variables*, Ann Arbor 1968.
- U.S. Department of Labor, Bureau of Labor Statistics, Bull. No. 1204, *Tables of Working Life for Women, 1950*, Washington 1957.
- Department of Labor, *Labour Gazette*, March, Ottawa 1966.

Production Scheduling, Intermediate Goods, and Labor Productivity

By W. H. LOCKE ANDERSON*

The empirical literature on the short-run relationship between labor input and output is at variance with the static neo-classical production model in two respects.¹ Measured elasticities of input with respect to output are less than one, and input seems to respond to output with a lag. The first of these empirical results seems inconsistent with the law of variable proportions; the second implies that firms operate below their short-run production functions for some or all of the time.

The various authors who have reported these statistical results have offered *ad hoc* rationalizations of varying degrees of rigor and explicitness of assumption. Although specific papers differ with respect to the emphasis which they place on particular elements in the cost picture, nearly all assert that adjustment costs are the key to understanding the differences between observation and the predictions of the static neoclassical model.

There seems to be little doubt that changes in labor input in either direction are costly. Ronald Soligo (pp. 174-75) offers a substantial list of costs to be attached to changes in employment, and Ray Fair (pp. 15-16) asserts with some persuasiveness that they apply as well to changes in the length of the work week. When adjustment costs are important, firms will operate on the static production

function only at seasonal or cyclical peaks. At other times they will "hoard" labor in order to avoid adjustment expenses. As a consequence, the proportional amplitude of fluctuations in labor input will be smaller than that of output, even though it might be larger if firms always operated on the static production function.² Moreover, if firms are uncertain about the sustainability of variations in their production levels, they will absorb such changes by varying their labor hoards until they become convinced that changes are to be lasting. Uncertainty will therefore introduce a lag into the relationship between production and labor input.

There is little reason to question the validity of this analysis. Adjustment costs, both monetary and imputed, clearly do influence the behavior of the firm in countless respects, and in the short run they may be as important as steady-state costs. The positions in which we observe firms are influenced not only by where they would like to be if they were at rest, but also by where they have recently been and where they plan to go.

Nonetheless, there is a strong possibility that the adjustment cost argument has been overstated. At least some portion of the apparent low elasticity and lagged adjustment may be attributed to the way in which the relationship between labor input and production is usually measured. If a firm produces intermediate goods which are storeable inputs into the production of its final product, the optimum

* The author is professor of economics at The University of Michigan. He would like to thank Barry Bosworth, Harold Shapiro, and Hugo Sonnenschein for their many helpful comments

¹ For an up to date bibliography, see Brechling and O'Brien. An excellent paper not listed in their bibliography is that of Fair.

² This is demonstrated by R. Solow, using variational methods and continuous time.

production schedules for these goods will often show smaller variations than those of the final products. Furthermore, the peaks and troughs in production of intermediate goods will often lag behind those in final products. If a substantial portion of the firm's labor force is engaged in the production of intermediate goods, any study which correlates total labor input to the output of final product alone is apt to indicate a low elasticity and a distributed lag even when the true elasticity is higher than unity and there is no lag.

To explore the cost relationships which lead to this state of affairs, the central portion of this paper is devoted to the production scheduling problem of a firm which produces a final good and an intermediate good under rising marginal costs with zero adjustment costs. After the optimum schedules have been derived, the total labor requirements are compared with production of the final good to reveal the extent to which this comparison leads to incorrect conclusions about the nature of the short-run input output relationship. In order that the analysis be relatively simple, the firm is assumed to have known deliveries, a fixed capital stock, and fixed wages and storage costs. Thus the model is more applicable to the analysis of seasonal variations in input and output than to cyclical variations.

I. *Scheduling a Single Product*

The exposition of production scheduling for firms with intermediate goods will be greatly simplified by solving first the scheduling problem for firms which produce no intermediate goods. This problem has been extensively studied by Franco Modigliani and Franz Hohn, Kenneth S. Arrow and S. Karlin, and Vernon Smith. This section adds little to their analysis. Its purpose is merely to apply their findings to the context within which we wish to study the two-good firm.

We begin by defining the notation and making explicit the assumptions. Let:

θ be an angle such that $2\pi/\theta = 1$ year;

$[0, T]$ = planning period (as the solution will show, this need not be much more than a year);

$S(t)$ = cumulative shipments from 0 to t ;

$X(t)$ = cumulative production from 0 to t ;

$H(t)$ = stocks at t ;

$L(t)$ = labor cost of production at t (assumed to be proportional to employment);

v = unit storage cost of output per unit (this is out-of-pocket cost, not interest).

The interest rate is taken to be zero. The time derivatives of S , X , and H can be interpreted as follows:

$\dot{S}(t)$ = level of shipments at t ;

$\dot{X}(t)$ = level of production at t ;

$\dot{H}(t)$ = rate of inventory accumulation at t where $\dot{H}(t) = \dot{X}(t) - \dot{S}(t)$.

Assume:

$H(0) = 0$;

$S(t) = \alpha t + (\beta/\theta) \sin t\theta$, and hence

$\dot{S}(t) = \alpha + \beta \cos t\theta$; $\alpha > \beta$ and hence $\dot{S}(t) > 0$;

$L(t) = a\dot{X}(t) + (c/2)\dot{X}(t)^2$.

No distinction is drawn between direct time and overtime. The right-hand side of the labor cost function is proportional to man-hours employed, the proportionality factor being an assumed constant wage rate. The constraints on production are:

$H(t) = X(t) - S(t) \geq 0$;

$\dot{X}(t) \geq 0$ (In the problem stated this constraint does not bind.)

The problem is to minimize the integral:

$$(1) \quad C = \int_0^T (vH(t) + L(t))dt,$$

which is the sum of storage and labor costs, subject to the nonnegativity constraints on stocks and production.

If we disregard the nonbinding constraint on $\dot{X}(t)$, we can see that the optimization problem is a common problem in the calculus of variations, that of minimizing the integral of a bounded function.³ The solution for $X(t)$, $t \in [0, T]$, must consist of portions which are either on the boundary $X(t) = S(t)$, or are extremals, that is, solutions to the Euler equation derived from (1). In principle, the solution may be entirely on the boundary or it may be entirely extremal.

Given the definitions of $H(t)$ and $L(t)$, the Euler equation for the minimization of C may be seen to be:

$$(2) \quad \dot{X}(t) = \frac{v}{c}.$$

Hence the extremal paths for $\dot{X}(t)$ and $X(t)$ are given by:

$$(3) \quad \dot{X}(t) = \frac{v}{c}t + k_1 \quad \text{and}$$

$$(4) \quad X(t) = \frac{v}{2c}t^2 + k_1t + k_2,$$

where k_1 and k_2 are constants of integration. The optimal level of $X(t)$ at any $t \in [0, T]$ will therefore either lie on the boundary $X(t) = S(t)$ or be given by the extremal (4).

To find a complete solution to the problem, it is necessary to identify those intervals over which $X(t)$ is extremal and those over which it is bounded. Let us

³ Elsgolc is a splendid introduction to the calculus of variations. The material necessary for the understanding of this paper is lucidly presented on pp. 13-39, 64-72, 94-98. The Appendix to this article sketches out an intuitive derivation of the solutions to both the one-good and the two-good problem without explicit use of variational methods.

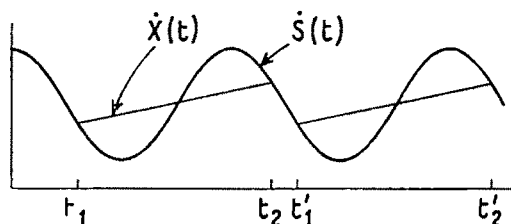


FIGURE 1

denote by t_1 and t_2 the entering and leaving points, respectively, of a segment of the optimal path which is extremal. Some necessary conditions must be met at these end points.⁴ They are:

$$(5) \quad \begin{aligned} X(t_1) &= S(t_1), \\ X(t_2) &= S(t_2), \\ \dot{X}(t_1) &= \dot{S}(t_1), \\ \dot{X}(t_2) &= \dot{S}(t_2). \end{aligned}$$

The first two simply state that net inventory accumulation over the entire period $t \in [t_1, t_2]$ must be zero; the second two state that $\dot{X}(t)$ must be continuous over $t \in [0, T]$.

These end-point conditions make it possible to visualize the entire solution with the aid of a diagram (Fig. 1). The sinusoid in Figure 1 is the level of shipments $\dot{S}(t)$. The straight line segments constitute those portions of the optimal path for which $\dot{X}(t)$ is extremal. Their slopes are equal to v/c , from (3). The end point conditions (5) enable us to find t_1 and t_2 explicitly from (3), (4), and the given form of $S(t)$, but the explicit solution is neither neat nor instructive.

Three interesting properties of the general nature of the solution may be seen from Figure 1.

1) If the seasonal pattern in shipments repeats itself over and over, so will the seasonal pattern in production. Thus t'_1 and t'_2 , the entering and leaving points of the second extremal portion occur one

⁴ See Elsgolc, pp. 94-98.

year later than the corresponding t_1 and t_2 .

2) The end of the planning horizon (T), as Modigliani and Hohn show, need not be much beyond the next seasonal peak in shipments.

3) Since the slopes of the extremal portions are v/c , the smaller is v (unit carrying cost) relative to c (the derivative of marginal cost with respect to production), the longer will be the interval $[t_1, t_2]$ and the larger will be the maximum inventory (achieved at 3 quarters of the way through each year). At one extreme when $v=0$, optimum production will be completely smooth from $t=t_1=\pi/2\theta$ onwards. At the other extreme, for $v/c \geq \max [S(t)] = \beta\theta$, there will be no storage and $X(t)=S(t)$ always. Note that for $\beta\theta > v/c \geq 2\beta\theta/\pi$, the amplitude of the seasonal fluctuations in production will be as great as that of shipments even though some storage takes place.

Extensions of this simple model may be made, but none need concern us at any great length here.

1) Initial stocks need not have been assumed to be zero. They only affect the optimal plan in the first period or so (and may require explicit recognition of the nonnegativity constraint in $\dot{X}(t)$). The effects of initial stocks are derived by Modigliani and Hohn, and by Arrow and Karlin.

2) The assumption of sinusoidal shipments was not essential. Any seasonal pattern with a single peak and a single trough would lead to a qualitatively similar solution.

3) Minor departures from the assumed quadratic labor cost function would change the shape of the extremal portions of the optimal path, but would not change the general outlines of the solution. For a general labor cost function, the Euler equation is:

$$(6) \quad \ddot{X}(t) = \frac{v}{\frac{dMC(\dot{X}(t))}{d\dot{X}(t)}},$$

where $MC(\dot{X}(t))$ is marginal labor cost as a function of output ($\dot{X}(t)$). Thus, if the marginal cost function is convex for a given output at a given time, the extremal passing through this point is concave, and if marginal cost is concave, the extremal is convex.

4) The introduction of a non-zero rate of interest makes little difference in the shape of the extremal paths unless the interest rate is large. If we seek to minimize discounted costs, we rewrite (1) as:

$$(1') \quad C = \int_0^T e^{-rt}(vH(t) + L(t))dt,$$

for which the Euler equation is:

$$(2') \quad \dot{X}(t) - r\dot{X}(t) = \frac{v - ra}{c}.$$

Its solution for $\dot{X}(t)$ is:

$$(3') \quad \dot{X}(t) = -\frac{v + ra}{rc} + ke^{rt},$$

where k is a constant of integration. The closeness of this solution to (3) may be seen if we compare the corresponding extremals through any point $(t_0, \dot{X}(t_0))$. From (3) we get

$$(7) \quad \dot{X}(t - t_0) = \dot{X}(t_0) + \frac{v}{c}(t - t_0).$$

Expanding e^{rt} in (3') and keeping only the linear term, (3') may be written as:

$$(7') \quad \dot{X}(t - t_0) = \dot{X}(t_0) + \frac{v + ra}{c}(t - t_0).$$

Thus with a positive interest rate the extremal is steeper than with a zero rate, a conclusion which would be reinforced by retaining more terms in the expansion. As

we should expect, therefore, the optimal plan with a positive interest rate involves more time on the boundary and smaller average stock holdings.

5) A storage capacity limitation which prevents the path shown in Figure 1 from being realized would lead to the solution shown in Figure 2, where for $t \in [t_2, t_3]$ stocks are at the permissible maximum.

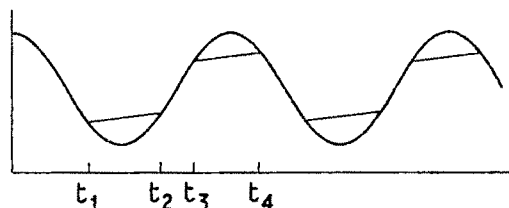


FIGURE 2

II. *Scheduling Intermediate and Final Products*

We turn now to a discussion of the optimal production plan when the firm produces an intermediate good which is an input into the production of the finished good. Define the units of the intermediate good so that one is required in the production of each unit of the finished good. Let:

θ again be an angle such that $2\pi/\theta = 1$ year;

$[0, T]$ = planning period;

$S(t)$ = cumulative shipments of the finished good from 0 to t ;

$X_1(t)$ = cumulative production of the finished good;

$X_2(t)$ = cumulative production of the intermediate good;

$H_1(t)$ = stocks of the finished good;

$H_2(t)$ = stocks of the intermediate good;

$L_1(t)$ = labor cost of production of the finished good at t ;

$L_2(t)$ = labor cost of production of the intermediate good at t ;

v_1 = unit storage cost of the finished good;

v_2 = unit storage cost of the intermediate good.

The time derivatives of S , X_1 , X_2 , H_1 , and H_2 have the obvious interpretations. Assume:

$$H_1(0), H_2(0) = 0;$$

$S(t)$ is given by the same function used

above;

$$L_1(t) = a_1 X_1(t) + \frac{c_1}{2} (X_1(t))^2;$$

$$L_2(t) = a_2 X_2(t) + \frac{c_2}{2} (X_2(t))^2.$$

The constraints are:

$$H_1(t) = X_1(t) - S(t) \geq 0;$$

$$H_2(t) = X_2(t) - X_1(t) \geq 0;$$

$\dot{X}_1(t), \dot{X}_2(t) \geq 0$ (again, these constraints will not bind).

The optimization problem is to minimize:

$$(8) \quad C = \int_0^T (v_1 H_1(t) + v_2 H_2(t) + L_1(t) + L_2(t)) dt,$$

subject to the nonnegativity constraints on stocks and production. Since the constraint on $H_1(t)$ is a given function of time, its effects on the optimal path of $S_1(t)$ may be found by application of end-point conditions (5). The constraint on $H_2(t)$ is not so simply handled, however, and it is convenient to recognize it by introducing a Lagrangian multiplier, $\lambda(t)$, where:

$$(9) \quad \begin{aligned} \lambda(t) &\equiv 0 && \text{whenever } H_2(t) > 0, \\ \lambda(t) &\neq 0 && \text{whenever } H_2(t) = 0. \end{aligned}$$

We then minimize the functional:

$$(10) \quad C^* = \int_0^T (v_1 H_1(t) + v_2 H_2(t) + L_1(t) + L_2(t) + \lambda(t) H_2(t)) dt.$$

The Euler equations are:

$$(11) \quad \dot{X}_1(t) = \frac{v_1 - v_2}{c_1} - \frac{\lambda(t)}{c_1}, \quad \text{and}$$

$$(12) \quad \dot{X}_2(t) = \frac{v_2}{c_2} + \frac{\lambda(t)}{c_2}.$$

Therefore, when $H_2(t)=0$, the extremal for $X_1(t)$ is given by the solution to:

$$(11') \quad \dot{X}_1(t) = \frac{v_1}{c_1 + c_2}.$$

But

$$(12') \quad \dot{X}_2(t) = \dot{X}_1(t) \quad (\text{i.e., } X_2(t) \text{ is under constraint}).$$

When $H_2(t) \neq 0$, the two extremals are given by solutions to:

$$(11'') \quad \dot{X}_1(t) = \frac{v_1 - v_2}{c_1}, \quad \text{and}$$

$$(12'') \quad \dot{X}_2(t) = \frac{v_2}{c_2}.$$

A comparison of (11') and (11'') is instructive. When $H_2(t)=0$, production of the finished good cannot diminish stocks of the intermediate good. The marginal storage cost is therefore v_1 and the derivative of marginal labor cost with respect to $\dot{X}_1(t)$ is $c_1 + c_2$, since additional production of the finished good must be matched by additional production of the intermediate good. When $H_2(t) > 0$, however, the net storage cost is only $v_1 - v_2$, since production of the finished good depletes stocks of the intermediate good. This being the case, the derivative of marginal labor cost with respect to $\dot{X}_1(t)$ is simply c_1 , since $\dot{X}_2(t)$ is not constrained to equal $\dot{X}_1(t)$.

We turn now to the question of determining which portions of the optimal paths are extremals and which lie on the boundaries imposed by the nonnegativity of stocks. If the optimum production schedule for the finished good is partly extre-

mal and partly bounded, then the time-path of $\dot{X}_1(t)$ is similar to the solution for $\dot{X}(t)$ given in Figure 1. It is clear that if $v_2/c_2 > (v_1 - v_2)/c_1$, then no part of the optimal path for $\dot{X}_2(t)$ can be extremal, since it is impossible to construct a path satisfying the end point conditions. Under these circumstances, the intermediate good will never be stocked, and the extremal for $\dot{X}_1(t)$ will have slope $v_2/(c_1 + c_2)$. Note that $v_2/c_2 \geq (v_1 - v_2)/c_1$ implies $v_2/c_2 \geq v_1/(c_1 + c_2)$. Bearing this in mind, we may distinguish five classes of optimal solution.

$$\text{Case 1:} \quad \frac{v_2}{c_2} < \frac{v_1 - v_2}{c_1} < \beta\theta.$$

Under these circumstances, the optimal plan (shown in Figure 3) involves accumulation of stocks of both goods during part of the year.

$$\text{Case 2:} \quad \frac{v_2}{c_2} < \beta\theta \leq \frac{v_1 - v_2}{c_1}.$$

The optimal path entails no stocking of the finished good, but the intermediate good is stocked over part of the year.

$$\text{Case 3:} \quad \beta\theta \leq \frac{v_2}{c_2} < \frac{v_1 - v_2}{c_1}.$$

Neither good is ever stocked.

$$\text{Case 4:} \quad \frac{v_2}{c_2} \geq \frac{v_1 - v_2}{c_1}.$$

$$\frac{v_1}{c_1 + c_2} < \beta\theta.$$

The intermediate good is never stocked but the final good is stocked over part of the year.

$$\text{Case 5:} \quad \frac{v_2}{c_2} \geq \frac{v_1 - v_2}{c_1}.$$

$$\frac{v_1}{c_1 + c_2} \geq \beta\theta.$$

Neither good is ever stocked.

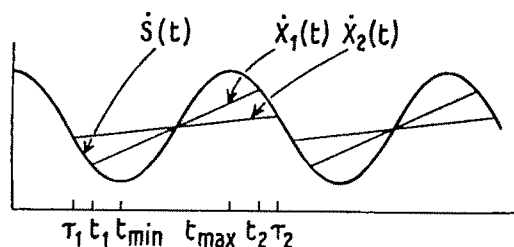


FIGURE 3

III. Actual and Apparent Labor Productivity

In order to discuss the implications of each of these patterns for productivity measurement, it will be convenient to plot the relationship between $\dot{X}_1(t)$ and $\dot{X}_2(t)$ on the $\dot{X}_1(t)$, $\dot{X}_2(t)$ phase plane. We introduce the following notation:

- t_1 = entering date of extremal portion of path of $\dot{X}_1(t)$;
- t_2 = leaving date of extremal portion of path of $\dot{X}_1(t)$;
- τ_1 = entering date of extremal portion of path of $\dot{X}_2(t)$;
- τ_2 = leaving date of extremal portion of path of $\dot{X}_2(t)$;
- t_{\min} = date of seasonal trough in $\dot{S}(t)$;
- t_{\max} = date of seasonal peak in $\dot{S}(t)$;

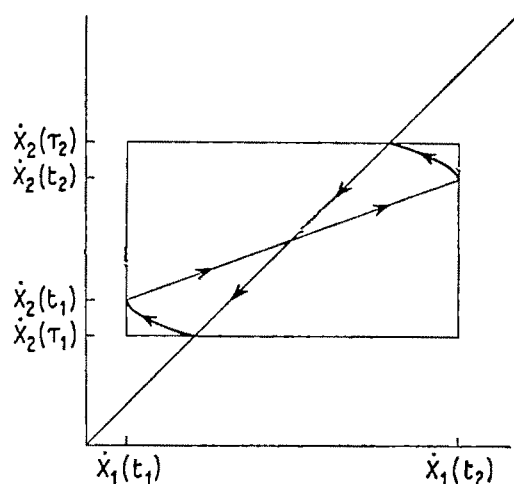


FIGURE 4

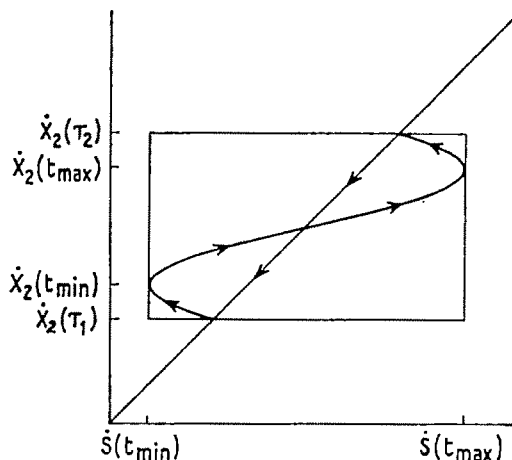


FIGURE 5

These dates are indicated in Figure 3.

Corresponding to Figure 3, we have the $\dot{X}_1(t)$, $\dot{X}_2(t)$ phase relationship shown in Figure 4, which depicts the solution to Case 1. The solution for Case 2, in which the optimal path for $\dot{X}_1(t)$ has no extremal portions, is shown in Figure 5. Since the optimal solutions for Cases 3, 4, and 5 always entail equality between $\dot{X}_1(t)$ and $\dot{X}_2(t)$, their phase paths (which are not illustrated) are simply 45 degree lines with amplitude depending on whether or not the finished good is stocked.

We proceed now to examine the seasonal variations in labor costs which accompany these seasonal variations in production of the two commodities. We may combine the two labor cost functions to get an expression for total labor cost.

$$L(t) = a_1 X_1(t) + \frac{c_1}{2} (X_1(t))^2 + a_2 X_2(t) + \frac{c_2}{2} (X_2(t))^2. \quad (13)$$

We may also define a "steady state" function in which $\dot{X}_1(t) = \dot{X}_2(t)$.

$$L^*(t) = (a_1 + a_2) \dot{X}_1(t) + \frac{c_1 + c_2}{2} (X_1(t))^2. \quad (14)$$

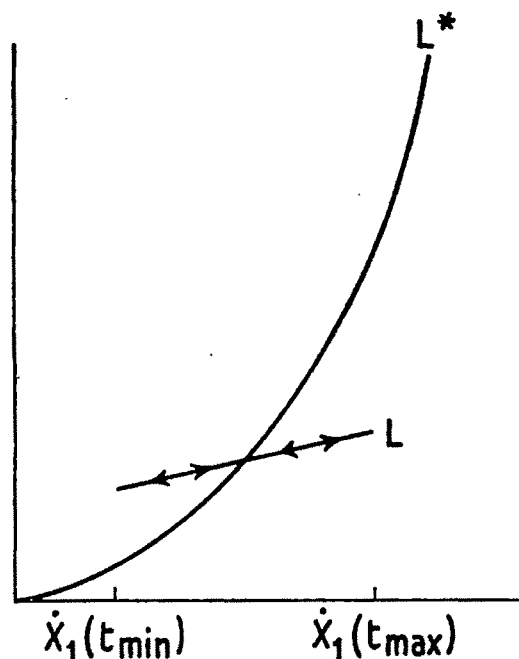


FIGURE 6

This shows the relationship between labor cost and production of the finished good when production of the intermediate good is in equilibrium with production of the finished good. If we wish to measure a cost function which uses only finished good production as its argument, it is presumably (14) which we seek to measure.

Using these relationships, we are in a position to see the various kinds of mischief which occur when we estimate cost functions without taking proper account of intermediate products.

Cases 3, 4, and 5 cause no difficulty; $\dot{X}_2(t)$ always equals $\dot{X}_1(t)$ and the L function always faithfully mirrors the L^* functions. For Cases 1 and 2, however, this is not so.

Consider first a degenerate version of Case 2, in which marginal production costs of $\dot{X}_1(t)$ are constant ($c_1=0$) and marginal storage costs of $\dot{X}_2(t)$ are zero ($v_2=0$). Under these circumstances the phase path is a horizontal line; $\dot{X}_1(t)$ will

never be stocked whereas $\dot{X}_2(t)$ will be produced at a constant rate (α) throughout the year. The true and apparent cost relations are shown in Figure 6. Although the true cost function is quadratic, with rising marginal cost and falling output per man-hour as output rises, the apparent cost function shows constant marginal cost and increasing output per man-hour. The apparent cost curve is only correct at the two times in the year at which $\dot{X}_1(t)$ and $\dot{X}_2(t)$ are equal.

An even more alarming distortion in the cost relationship occurs in Case 1 and the nondegenerate ($c_1, v_2 \neq 0$) version of Case 2. During the seasonal phase from τ_1 to the next year's τ_1 , $\dot{X}_2(t)$ and $\dot{X}_1(t)$ are equal and a portion of the L^* function is traced out. Subsequently, however, $\dot{X}_2(t)$ lies above $\dot{X}_1(t)$ during the seasonal trough in $\dot{X}_1(t)$ and below it during its seasonal peak (see Figures 4 and 5). The effect of this on the apparent cost curve may be seen in Figure 7, which corresponds to the Case 1 shown in Figure 4. The path of L , which is traced out by the arrowed phase line, is for much of the year considerably

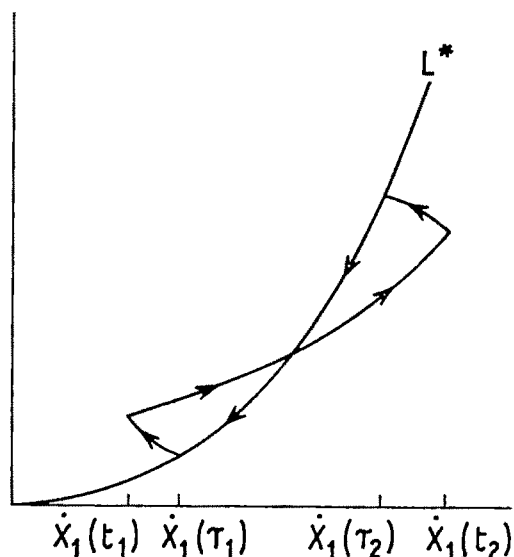


FIGURE 7

Multiperiod Consumption-Investment Decisions

By EUGENE F. FAMA*

I. The Problem

The simplest version of the multiperiod consumption-investment problem considers a consumer with wealth w_1 , defined as the market value of his assets at the beginning of period 1, which must be allocated to consumption c_1 and a portfolio investment $w_1 - c_1$. The portfolio will yield an uncertain wealth level w_2 at the beginning of period 2 which must be divided between consumption c_2 and investment $w_2 - c_2$. Consumption-investment decisions must be made at the beginning of each period, until the consumer dies and his wealth is distributed among his heirs. The consumer's objective is to maximize the expected utility of lifetime consumption.

Uncertainty models of the multiperiod consumption-investment problem have been considered by Edmund Phelps, Nils Hakansson, and Jan Mossin (1968). But their quite similar treatments place severe restrictions on both the form of the consumer's utility function and the process generating investment returns. For example, the most general model is Hakansson's. He assumes that the probability distributions of one-period portfolio wealth relatives¹ that will be available at any future period t are known for certain at

period 1 and thus are independent of events that will occur between periods 1 and t . The consumer's utility function is assumed to be of the additive form

$$u(c_1, \dots, c_t, \dots) = \sum_{t=1}^{\infty} \alpha^{t-1} U(c_t), \\ 0 < \alpha < 1,$$

so that the utility provided by consumption in period t cannot be affected by levels of consumption attained in other periods. Moreover, the one-period utility function $U(c_t)$ is assumed to be monotone increasing and strictly concave (i.e., marginal utility is positive and the consumer is a risk averter),² and $U(c_t)$ must imply either "constant risk aversion" or "constant proportional risk aversion," where these terms are as defined by John Pratt.

Hakansson shows that in this model the optimal consumption for any period t is a linear increasing function of wealth w_t . In addition, the portfolio opportunities that will be available in periods after t affect the optimal split of w_t between current consumption and investment, but the optimal *proportions* of portfolio funds invested in different assets at t depend only on the consumer's one-period utility function $U(c_t)$ and on the distributions of one-period wealth relatives associated with currently available portfolios. In essence, the choice of an optimal portfolio *mix* is "myopic" in the sense that it depends only on one-period utilities and returns.

But these appealingly simple results are direct consequences of the restrictions

* Graduate School of Business, University of Chicago. The author wishes to thank Jacques Dréze, Gary Eppen, Michael Jensen, a referee, and especially Merton H. Miller for their most helpful comments. Work on this project was supported by a grant from the National Science Foundation.

¹ The one-period wealth relative from t to $t+1$ is defined as total dollars of market value at $t+1$ per dollar invested at t . It is thus one plus the one-period (percentage) return from t to $t+1$.

² A general definition of concavity is provided in fn. 9.

imposed on utility functions and the process generating investment returns by the Phelps-Hakansson-Mossin models. The goal in this paper is to present a more general multiperiod consumption-investment model, but one which nevertheless leads to interesting hypotheses about observable aspects of consumer behavior. The main result is the proposition that if the consumer is risk averse (i.e., his utility function for *lifetime* consumption is strictly concave) and markets for consumption goods and portfolio assets are perfect,³ then the consumer's observable behavior in the market in any period is indistinguishable from that of a risk averse expected utility maximizer who has a one-period horizon.

With this result it is then possible to provide a *multiperiod* setting for hypotheses about consumer behavior derived from *one-period* wealth allocation models, and these have been studied extensively.⁴ One-period models assume, of course, that consumers have one-period horizons, but in most cases their behavioral propositions only require that consumers behave *as if* they are risk averse one-period expected utility maximizers, and this will be the case in the multiperiod model to be presented here. Thus perhaps the major contribution of this paper is in providing a means for bridging the gap between one-period and multiperiod models.

As a specific illustration we will later consider in detail the adjustments to the multiperiod model that are necessary to

provide a multiperiod setting for the major propositions about consumer behavior associated with the one-period, two-parameter wealth allocation models of Harry Markowitz, James Tobin (1958, 1965), William Sharpe (1963, 1964), John Lintner (1965a, 1965b), and Fama (1965, 1968a). Indeed it will be shown that a multiperiod model in which the optimal portfolio for any period is "efficient" in terms of distributions of *one-period* portfolio wealth relatives requires few assumptions beyond those already made in the one-period models.⁵

II. The Wealth Allocation Model

First the multiperiod model must itself be developed. Let β_t , the "state of the world," signify the set of events (current and past prices, etc.) that constitutes history up to t . Thus β_{t-1} is a subset of β_t . If the state of the world at t is β_t , there will be $n(\beta_t)$ investment assets available to the consumer; the wealth relatives from t to $t+1$ for these assets are represented by the vector

$$\begin{aligned} R &= R(\beta_{t+1}) \\ (1) \quad &= (r_1(\beta_{t+1}), r_2(\beta_{t+1}), \dots, \\ &\quad r_{n(\beta_t)}(\beta_{t+1})), \end{aligned}$$

so that a value of β_{t+1} implies a value of the vector of wealth relatives. If

$$H_{\beta_t} = (h_1, h_2, \dots, h_{n(\beta_t)})$$

is the (nonnegative) vector of dollars invested in each asset at t in state β_t , the consumer's wealth at $t+1$ will be

$$(2) \quad w_{t+1} = H_{\beta_t} R(\beta_{t+1})',$$

where w_{t+1} and $R(\beta_{t+1})$ are random vari-

³ That is: (a) consumption goods and portfolio assets are infinitely divisible, (b) reallocations of consumption and investment expenditures are costless, and (c) the consumer's activities in any market have a negligible effect on prices. Such a "perfect markets" assumption is common to almost all wealth allocation models (one-period and multiperiod) and will be maintained throughout the discussion here.

⁴ See, for example, Peter Diamond, Jacques Drèze and Franco Modigliani (1966), Fama (1965, 1968a, 1968b), Jack Hirshleifer (1965, 1966), Michael Jensen, John Lintner (1965a, 1965b).

⁵ Measuring the dispersion of a distribution of wealth relatives with a single parameter such as the standard deviation or semi-interquartile range, a portfolio is efficient in the Markowitz sense if no other portfolio with the same expected one-period wealth relative has lower dispersion, and no other portfolio with higher expected wealth relative has the same or lower dispersion.

ables at t , and the primed variable denotes the transpose.

The consumer's behavior is assumed to conform to the von Neumann-Morgenstern expected utility model. Thus if for simplicity we initially assume he will die for certain⁶ at the beginning of period $\tau+1$, and if the state of the world at $\tau+1$ is $\beta_{\tau+1}$, the consumer's utility for lifetime consumption is given by the "cardinal" function

$$U_{\tau+1}(C_{\tau+1} | \beta_{\tau+1}) \\ = U_{\tau+1}(c_{1-k}, \dots, c_1, \dots, c_{\tau+1} | \beta_{\tau+1}),$$

where in general

$$C_t = (c_{1-k}, \dots, c_1, \dots, c_t)$$

is consumption from the beginning of his life, period $1-k$, through period t , and the consumption $c_{\tau+1}$ is in the form of a bequest.⁷ The goal of the consumer in his consumption-investment decisions is to maximize the expected utility of *lifetime* consumption.

The consumer must make an optimal consumption-investment decision for period 1, taking into account that decisions must also be made at the beginning of each future period prior to $\tau+1$, and that these future decisions will depend on future events. Dynamic programming, with its "backward optimization," provides a natural approach. That is, to solve the decision problem for period 1, the consumer first determines optimal decisions for all contingencies for the decision problem to

be faced at period τ . Then he determines optimal decisions for $\tau-1$, under the assumption that he will always make optimal decisions at τ . And so on, until he works his way back to the decision at period 1, which is then based on the assumption that optimal decisions will be made at each future period for any possible contingency.

Formally, optimal decisions for all w_t and β_t can be summarized by the function

$$U_{\tau}(C_{\tau-1}, w_{\tau} | \beta_{\tau}) \\ (3) \quad = \max_{c_{\tau}, H_{\beta_{\tau}}} \int_{\beta_{\tau+1}} U_{\tau+1}(C_{\tau}, HR' | \beta_{\tau+1}) \cdot dF_{\beta_{\tau}}(\beta_{\tau+1}),$$

subject to the constraints

$$0 \leq c_{\tau} \leq w_{\tau}, \quad H_{\beta_{\tau}} i' = w_{\tau} - c_{\tau}, \\ H_{\beta_{\tau}} \geq 0_n(\beta_{\tau}),$$

where $F_{\beta_{\tau}}(\beta_{\tau+1})$ is the distribution function of $\beta_{\tau+1}$ given state β_{τ} at τ ; $0_n(\beta_{\tau})$ is the null vector (i.e., a vector of zeros) with dimension $1 \times n(\beta_{\tau})$; i is the sum vector (i.e., a vector of ones) which will always be assumed to have whatever dimension is needed for the purpose at hand.⁸ The function $U_{\tau}(C_{\tau-1}, w_{\tau} | \beta_{\tau})$ is the maximum of expected utility at τ as a function of realized past consumption $C_{\tau-1}$ and current wealth w_{τ} , given that the state of the world is β_{τ} .

More generally, for $t=1, 2, \dots, \tau$, the process of backward optimization is summarized by the recursive relation

$$(4) \quad U_t(C_{t-1}, w_t | \beta_t) = \\ \max_{c_t, H_{\beta_t}} \int_{\beta_{t+1}} U_{t+1}(C_t, HR' | \beta_{t+1}) \cdot dF_{\beta_t}(\beta_{t+1}),$$

subject to

$$0 \leq c_t \leq w_t, \quad H_{\beta_t} i' = w_t - c_t, \\ H_{\beta_t} \geq 0_n(\beta_t).$$

⁸ Keep in mind that, given (2), integrating over $\beta_{\tau+1}$ implies averaging over $R=R(\beta_{\tau+1})$ and thus over $w_{\tau+1}=H_{\beta_{\tau}}R'$.

⁶ Later the model will be extended to allow for an uncertain period of death, (Section IV).

⁷ Since the utility of a given $C_{\tau+1}$ can depend on the state of the world $\beta_{\tau+1}$, the model is consistent with the so-called "state preference" model introduced by Kenneth Arrow. As we shall see later, the more traditional framework in which utilities are not state dependent is just a special case.

For an axiomatic development of the expected utility model which implies the existence of subjective probabilities and allows for state dependent utilities (see Drèze). Drèze's analysis is in turn an extension of Savage's work.

The function $U_t(C_{t-1}, w_t | \beta_t)$ provides the maximum expected utility of *lifetime* consumption if the consumer is in state β_t at period t , his wealth is w_t , his past consumption was C_{t-1} , and optimal consumption-investment decisions are made at the beginning of period t and all future periods.

Expression (4) exemplifies a common feature of dynamic programming models. In general it is possible to *represent* the decision problem of any period t in terms of a derived objective function (in this case U_{t+1}) which is *explicitly* a function only of variables for $t+1$ and earlier periods, but which in fact summarizes the results of optimal decisions at $t+1$ and subsequent periods for all possible future events. Thus the recursive relation (4) *represents* the multiperiod problem as a sequence of "one-period" problems, though at any stage in the process the objective function used to solve the one-period problem summarizes optimal decisions for all future periods.

Representing the multiperiod consumption-investment problem as a sequence of one-period problems in itself says nothing about the characteristics of an optimal decision for any period. The main result of this paper is, however, the following.

Proposition 1. If the utility function for lifetime consumption $U_{t+1}(C_{t+1} | \beta_{t+1})$ has properties characteristic of risk aversion (specifically, if for all β_{t+1} , $U_{t+1}(C_{t+1} | \beta_{t+1})$ is monotone increasing and strictly concave in C_{t+1}), then for all t the derived functions $U_t(C_{t-1}, w_t | \beta_t)$ will also have these properties.⁹

⁹ The monotonicity of U_{t+1} says that the marginal utility of consumption in any period is positive, while strict concavity implies that for $0 < \alpha < 1$,

$$U_{t+1}(\alpha C_{t+1} + (1 - \alpha)\hat{C}_{t+1} | \beta_{t+1}) > \alpha U_{t+1}(C_{t+1} | \beta_{t+1}) + (1 - \alpha)U_{t+1}(\hat{C}_{t+1} | \beta_{t+1}),$$

where C_{t+1} and \hat{C}_{t+1} are any two consumption vectors that differ in at least one element. Geometrically, concavity says that a straight line between any two points on the function U_{t+1} lies below the function. As in the

The proof of the proposition is presented in the Appendix.

III. Implications

Though at this point its importance is far from obvious, it is the concavity of the functions $U_t(C_{t-1}, w_t | \beta_t)$ for all t and β_t , as stated in Proposition 1, that will now allow us to bridge the gap between one-period and multiperiod wealth allocation models.

A. The Utility of Money Function

A foretaste of the discussion can be obtained by using the multiperiod model to derive the familiar utility of money function, most often discussed in the literature in connection with the expected utility model. If the state of the world at period 1 is β_1 and the consumer's past consumption has been \hat{C}_0 , then for $t=1$ expression (4) yields

$$\begin{aligned} v_1(w_1 | \beta_1) &= U_1(\hat{C}_0, w_1 | \beta_1) \\ &= \max_{c_1, H_{\beta_1}} \int_{\beta_2} U_2(C_1, HR' | \beta_2) dF_{\beta_1}(\beta_2), \end{aligned}$$

subject to

$$\begin{aligned} 0 &\leq c_1 \leq w_1, \quad H_{\beta_1} c_1' = w_1 - c_1, \\ H_{\beta_1} &\geq 0_{\pi(\beta_1)}. \end{aligned}$$

v_1 is the relevant utility function for timeless gambles taking place at period 1: that is, gambles where the outcome is known before the consumption-investment decision of period 1 is made. From Proposition 1, v_1 has the characteristics of a risk averter's utility of money function: that is, it is monotone increasing and strictly concave in w_1 . Thus, though he obtains his utility of money function by a complicated process of backward optimization, and though his utility of money function in fact shows the expected utility of lifetime consumption associated with a given level of wealth at period 1, the consumer's

case of the more familiar utility of money function, the concavity of U_{t+1} implies risk aversion.

behavior in choosing among timeless gambles is indistinguishable from that of a risk averter making a once-and-for-all decision. In other words, our analysis provides a multiperiod setting for the more traditional discussions of utility of money functions for risk averters, most of which abstract from the effects of future decisions.

B. One-Period and Multiperiod Models: General Treatment

More generally, when it comes time to make a decision at the beginning of any period t , $t=1, 2, \dots, \tau$, past consumption (equal, say, to \hat{C}_{t-1}) is known, so that the decision at t can be based on the function

$$v_{t+1}(c_t, w_{t+1} | \beta_{t+1}) = U_{t+1}(\hat{C}_{t-1}, c_t, w_{t+1} | \beta_{t+1})$$

Thus, for given wealth w_t and state of the world β_t , the consumer's problem at t can be expressed as

$$(5) \quad \max_{c_t, H_{\beta_t}} \int_{\beta_{t+1}} v_{t+1}(c_t, HR' | \beta_{t+1}) dF_{\beta_t}(\beta_{t+1}),$$

subject to

$$0 \leq c_t \leq w_t, \quad H_{\beta_t} i' = w_t - c_t,$$

$$H_{\beta_t} \geq 0_{\pi(\beta_t)}.$$

Since, from Proposition 1, U_{t+1} is monotone increasing and strictly concave in (C_t, w_{t+1}) , v_{t+1} is monotone increasing and strictly concave in (c_t, w_{t+1}) . Thus, though the consumer faces a τ period decision problem, the function $v_{t+1}(c_t, w_{t+1} | \beta_{t+1})$, which is relevant for the consumption-investment decision of period t , has the properties of a risk averter's *one-period* utility of consumption-terminal wealth function. Though the consumer must solve a multiperiod problem, given v_{t+1} his observed behavior in the market is indistinguishable from that of a risk averse expected utility maximizer who has a one-period horizon.¹⁰

¹⁰ But we must keep in mind that though $v_{t+1}(c_t, w_{t+1} | \beta_{t+1})$ is only explicitly a function of variables for

In itself, this result says little about consumer behavior. Its value derives from the fact that it can be used to provide a multiperiod setting for more detailed behavioral hypotheses usually obtained from specific one-period model. Since by design the multiperiod model is based on less restrictive assumptions than most one-period models,¹¹ adapting it to any specific one-period model will require additional assumptions. But as we shall now see, these are mostly restrictions already implicit or explicit in the one-period models. Little generality is lost in going from a one-period to a multiperiod framework.

C. A Multiperiod Setting for One-Period, "Two-Parameter" Portfolio Models

Given the concavity of v_{t+1} , (5) is formally equivalent to the consumption-investment problem of a risk averse consumer with state dependent utilities and a one-period horizon.¹² As such it can be used to provide a multiperiod setting for a wide

periods t and $t+1$, it shows the maximum expected utility of *lifetime* consumption, given optimal consumption-investment decisions in periods subsequent to t . Thus v_{t+1} depends both on tastes, as expressed by the function $U_{t+1}(C_{t+1} | \beta_{t+1})$, and on the consumption-investment opportunities that will be available in future periods.

The notion of summarizing market opportunities in a utility function should not cause concern. Indeed this is done when utility is written (as we have done throughout) as a function of consumption dollars; then we are implicitly summarizing the consumption opportunities (in terms of goods and services and their anticipated prices) that will be available in each period. We shall return to this point in the Appendix where the utility function $U_{t+1}(C_{t+1} | \beta_{t+1})$ for dollars of consumption will be derived from a more basic utility function for consumption goods.

¹¹ In particular, we have essentially assumed only that markets for consumption goods and portfolio assets are perfect, and that the consumer is a risk averter in the sense that his utility function for lifetime consumption is strictly concave.

¹² The term "state dependent utilities" refers to the fact that the function $v_{t+1}(c_t, w_{t+1} | \beta_{t+1})$ allows the utility of a given combination (c_t, w_{t+1}) to depend on β_{t+1} . Hirschleifer (1965, 1966) uses instead the term "state preference" to refer to this condition.

variety of one-period models such as, for example, the one-period model analyzed in detail by Hirschleifer.

But the theories of wealth allocation most thoroughly discussed in the literature are the one-period, two-parameter portfolio models of Markowitz, Tobin (1958, 1965), Sharpe (1963), and Fama (1965, 1968a). These have in turn been used by Sharpe (1964), Lintner (1965a, 1965b), Mossin (1966), and Fama (1968a, 1968b) as the basis of one-period theories of capital market equilibrium. The market equilibrium relationships between the one-period expected wealth relatives and risks of individual securities and portfolios derived from these models have in turn been given some empirical support by Marshall Blume and Michael Jensen. The remainder of this section will be concerned with using our model to provide a multiperiod setting for the apparently useful results of these one-period models.

The two-parameter portfolio models start with the assumption that one-period wealth relatives on assets and portfolios conform to two-parameter distributions of the same general type. That is, the distribution for any asset or portfolio can be fully described once its expected value and a dispersion parameter, such as the standard deviation or the semi-interquartile range, are known.¹³ It is then shown that if

¹³ Since assets and portfolios must have distributions of the same two-parameter "type," the analysis is limited to the class of symmetric stable distributions, which includes the normal as a special case. Properties of these distributions are discussed, for example, in Fama (1965) and Benoit Mandelbrot, and a discussion of their role in portfolio theory can be found in Fama (1968a).

Alternatively, the results of the mean-standard deviation version of the two-parameter portfolio models can be obtained by assuming that one-period utility functions are quadratic in w_{t+1} . But strictly speaking, since the quadratic implies negative marginal utility at high levels of w_{t+1} , it is not a legitimate utility function. Moreover, the empirical evidence (see Marshall Blume, Fama (1965), Mandelbrot, and Richard Roll) that distributions of security and portfolio wealth relatives conform well to the infinite variance members of the

investors behave as if they try to maximize expected utility with respect to one-period utility functions $v_{t+1}(c_t, w_{t+1})$ that are strictly concave in (c_t, w_{t+1}) , optimal portfolios will be efficient in terms of the two parameters of distributions of one-period wealth relatives.¹⁴ The fact that optimal portfolios must be efficient then makes it possible to derive market equilibrium relations between expected wealth relatives and measures of risk for individual assets and portfolios.

But these models assume somewhat more about the utility function v_{t+1} than our multiperiod model. In particular, in the multiperiod model the function $v_{t+1}(c_t, w_{t+1}|\beta_{t+1})$, which is relevant for the consumption-investment decision of period t , is strictly concave in (c_t, w_{t+1}) , but utility can be a function of the state β_{t+1} (i.e., utility can be state dependent). Thus to provide a multiperiod setting for the one-period two-parameter models it is sufficient to determine conditions under which v_{t+1} will be independent of β_{t+1} .

State dependent utilities in the derived functions v_{t+1} have three possible sources. First, tastes for given bundles of consumption goods can be state dependent. Second, as will be shown in Proposition 2 of the Appendix, utilities for given dollars of consumption depend on the available consumption goods and services and their prices, and these are elements of the state of the world. Finally, the investment opportunities available in any given future period may depend on events occurring in preceding periods, and such uncertainty about investment prospects induces state dependent utilities. Thus the most direct way to exclude state dependent utilities

symmetric stable class casts doubt on any model that relies on the existence of variances. Since the approach based on general two-parameter return distributions by-passes these problems, it seems simplest to lay the quadratic to rest, at least for the purpose of portfolio models.

¹⁴ This concept of efficiency was defined in fn. 5.

is to assume that the consumer behaves as if the consumption opportunities (in terms of goods and services and their prices) and the investment opportunities (distributions of one-period portfolio wealth relatives) that will be available in any future period can be taken as known and fixed at the beginning of any previous period, and that the consumer's tastes for given bundles of consumption goods and services are independent of the state of the world.¹⁵

With these assumptions, the utility of a given (c_t, w_{t+1}) is independent of β_{t+1} , so that β_{t+1} can be dropped from $v_{t+1}(c_t, w_{t+1} | \beta_{t+1})$. Thus for given wealth, w_t , the decision problem facing the consumer at the beginning of any period t can be written as

$$\max_{c_t, H\beta_t} \int_{R_{t+1}} v_{t+1}(c_t, HR') dF(R_{t+1})$$

Subject to

$$\begin{aligned} 0 \leq c_t \leq w_t, \quad H\beta_t \mathbf{i}' &= w_t - c_t, \\ H\beta_t &\geq 0_n(\beta_t), \end{aligned}$$

where $F(R_{t+1})$ is the distribution function for the vector of wealth relatives R_{t+1} .

Since Proposition 1 applies directly to this simplified version of the multiperiod model, at any period t the function $v_{t+1}(c_t, w_{t+1})$ is monotone increasing and strictly concave in (c_t, w_{t+1}) and is thus formally equivalent to the one-period utility function used in the standard treatments of the one-period, two-parameter portfolio models. If distributions of one-period security and portfolio wealth relatives are of the same two-parameter type, we have a multiperiod model in which the consumer's behavior each period is indistinguishable from that of the consumer in the traditional one-period, two-

parameter portfolio models. From here it is a short step to develop a multiperiod setting for period-by-period application of the major results of the one-period, two-parameter models of market equilibrium.

IV. Extensions: Uncertain Period of Death

For simplicity, the development of Proposition 1 and its implications made use of the simplest version of the multiperiod consumption-investment model. In particular, it was assumed that (a) the consumer's resources at the beginning of any period t consist entirely of w_t , the value of the marketable assets carried into the period from previous periods; and that (b) the period of the consumer's death is known. But it is not difficult (indeed the major complications are notational) to extend the model to take account of the fact that the consumer has an asset, his "human capital," which will generate income in periods subsequent to t , but which cannot be sold outright in the market. The extended model would allow the ways that the consumer employs his human capital during t —his choice of occupation (s) and the division of his time between labor and leisure—to be at his discretion. It is also easy to extend the model to allow for opportunities the consumer may have to borrow against future labor income or against his portfolio.

But these extensions will not be pursued here.¹⁶ We shall consider instead how the possibility of an uncertain period of death can be introduced into the simple wealth allocation model of Section II.

For simplicity, the analysis so far has assumed that the consumer dies for certain at the beginning of period $\tau+1$. But the model, exactly as stated in (4), is consistent with the probabilistic occurrence of

¹⁵ It is important to note that some such assumptions are implicit in the one-period, two-parameter models themselves since they do not allow for the effects of state dependent utilities on the consumption-investment decision. Exactly these assumptions are quite explicit in the Phelps-Hakansson-Mossin models.

¹⁶ They are discussed in detail in Fama (1969), an earlier version of this paper, which will be made available to readers on request.

death (and the distribution of the consumer's wealth among his heirs) in earlier periods. A subset of the events that comprise the state of the world, β_t , is the set of all events up to t that could affect the consumer's utility for any vector of lifetime consumption. A subset of these events could in turn be the life-death status of the consumer. Thus the state of the world at t might be defined as

$$\beta_t = (z_t, \hat{\beta}_t),$$

where the variable z_t represents the life-death status of the consumer and can take either the value a_t (indicating that the consumer is alive at t), or the value d_t (indicating that death occurred in some period $\hat{t} \leq t$), and where $\hat{\beta}_t$ is the set of all other elements of the state of the world. With this interpretation of β_t , it is easy to see that the model presented above (specifically, in (4)) is consistent with the possibility of probabilistic occurrence of death in periods prior to $\tau+1$.

Nevertheless some interesting insights into the role of the horizon period $\tau+1$ can be obtained by examining the effects of "probabilistic death" in a little more detail. When the period of death is uncertain, the consumer must make an optimal consumption-investment decision for period 1, taking into account that decisions must also be made at the beginning of any future period at which he is alive, but that the decision process will terminate as soon as he dies.

If the consumer is alive at τ , optimal decisions for all w_τ and β_τ can be summarized by the function

$$(6) \quad U_\tau(C_{\tau-1}, w_\tau | \beta_\tau) = U_\tau(C_{\tau-1}, w_\tau | a_\tau, \hat{\beta}_\tau) \\ = \max_{c_\tau, H_{\beta_\tau}} \int_{\beta_{\tau+1}} U_{\tau+1}(C_\tau, HR' | d_{\tau+1}, \hat{\beta}_{\tau+1}) \\ \cdot dF_{\beta_\tau}(\beta_{\tau+1}),$$

subject to

$$0 \leq c_\tau \leq w_\tau, \quad H_{\beta_\tau} i' = w_\tau - c_\tau, \\ H_{\beta_\tau} \geq 0_n(\beta_\tau),$$

where in this case $\beta_\tau = (a_\tau, \hat{\beta}_\tau)$ and $\beta_{\tau+1} = (d_{\tau+1}, \hat{\beta}_{\tau+1})$. Expression (6) is just (4) when $t = \tau$ and the consumer is alive at τ .

On the other hand, if the consumer dies at the beginning of any period t ($t = 1, 2, \dots, \tau+1$), his wealth is immediately distributed among his heirs, and expression (4) for the expected utility of his lifetime consumption becomes

$$(7) \quad U_t(C_{t-1}, w_t | \beta_t) = U_t(C_{t-1}, w_t | d_t, \hat{\beta}_t) \\ = \int_{\beta_{\tau+1}} U_{\tau+1}(C_t, 0_{\tau+1-t} | \beta_{\tau+1}) dF_{\beta_t}(\beta_{\tau+1}),$$

where $w_t = c_t$ is his bequest, and in this case $\beta_t = (d_t, \hat{\beta}_t)$ and $\beta_{\tau+1} = (d_t, \hat{\beta}_{\tau+1})$.¹⁷

In his consumption-investment decision for any period prior to τ , the consumer must consider that he could be either alive or dead at the beginning of the following period. Assuming for simplicity that the occurrence of death is independent of other elements of the state of the world, let x_t be the conditional probability that the consumer will be alive at t , given that he is alive at $t-1$. Then, with (6) as a starting point, for $t = 1, 2, \dots, \tau-1$, the process of backward optimization summarized by (4) is now expressed by (7) and the recursive relations

$$(8) \quad U_t(C_{t-1}, w_t | z_t, \hat{\beta}_t) \\ = x_t U_t(C_{t-1}, w_t | a_t, \hat{\beta}_t) \\ + (1 - x_t) U_t(C_{t-1}, w_t | d_t, \hat{\beta}_t), \\ U_t(C_{t-1}, w_t | a_t, \hat{\beta}_t) \\ (9) \quad = \max_{c_t, H_{\beta_t}} \int_{\beta_{t+1}} U_{t+1}(C_t, HR' | z_{t+1}, \hat{\beta}_{t+1}) \\ \cdot dF_{\beta_t}(\beta_{t+1}),$$

subject to

$$0 \leq c_t \leq w_t, \quad H_{\beta_t} i' = w_t - c_t, \\ H_{\beta_t} \geq 0_n(\beta_t).$$

¹⁷ If the consumer is not concerned with events subsequent to his death, then

$U_t(C_{t-1}, w_t | d_t, \hat{\beta}_t) = U_{\tau+1}(C_t, 0_{\tau+1-t} | d_t, \hat{\beta}_{\tau+1})$ for $w_t = c_t$ and all $\hat{\beta}_{\tau+1}$ such that $\hat{\beta}_t$ is a subset of $\hat{\beta}_{\tau+1}$.

The function $U_t(C_{t-1}, w_t | a_t, \beta_t)$ in (9) provides the maximum expected utility of lifetime consumption and bequests if the consumer is alive in state β_t at period t , his wealth is w_t , his past consumption was C_{t-1} , and optimal consumption-investment decisions are made at the beginning of period t and all future periods at which he is alive.

Since (7)–(9) are just a special case of the model summarized by (4), Proposition 1 applies directly to the probabilistic death model. In this case the proposition implies that for all t and β_t , $U_t(C_{t-1}, w_t | z_t, \beta_t)$, $U_t(C_{t-1}, w_t | a_t, \beta_t)$, and $U_t(C_{t-1}, w_t | d_t, \beta_t)$ are monotone increasing and strictly concave in (C_{t-1}, w_t) .

Finally, expressions (7)–(9) suggest an alternative to the “sure death” interpretation of the horizon $\tau+1$. For given wealth w_1 , and state of the world β_1 , the consumer's problem at period $\tau+1$ is to choose c_1 and H_{β_1} which

$$\begin{aligned} & \max_{c_1, H_{\beta_1}} \int_{\beta_2} U_2(C_1, HR' | z_2, \beta_2) dF_{\beta_1}(\beta_2) \\ &= \max_{c_1, H_{\beta_1}} \int_{\beta_2} [x_2 U_2(C_1, HR' | a_2, \beta_2) \\ & \quad + (1 - x_2) U_2(C_1, HR' | d_2, \beta_2)] dF_{\beta_1}(\beta_2). \end{aligned}$$

Using (8) and (9) to expand this expression, it can be shown that the decision problem of period t has weight

$$x_t x_{t-1} \cdots x_2 = x(a_t | a_1)$$

in the expected utility for the decision of period 1. The probability $x(a_t | a_1)$ of being alive at t will decrease with t . Thus in general for some $t = \tau+1$, the decisions of periods $\tau+1$ and beyond will have negligible weight in the expected utility for the decision at period 1, so that in the decision at period 1 it is unnecessary to look beyond $\tau+1$.

More simply, since the consumer is likely to be dead, the effects of distant future decisions, which are unlikely to be made, can be ignored. And this result does

not arise from discounting of future consumption, though the effect is the same. At period 1, consumption in period $\tau+1$ may be regarded as equivalent to consumption in period 1. But in the decision of period 1 the decision of $\tau+1$ is weighted by the probability that the consumer will be alive at that time, which reduces the importance of the future consumption in the current decision.

V. Conclusion

In sum, assuming only that markets for consumption goods and portfolio assets are perfect and that the consumer is risk averse in the sense that his utility function for lifetime consumption is strictly concave, it has been shown that though he faces a multiperiod problem, in his consumption-investment decision for any period the consumer's behavior is indistinguishable from that of a risk averter who has a one-period horizon. It was then shown how this result can be used to provide a multiperiod setting for the more detailed hypotheses about risk averse consumer behavior that are traditionally derived in a one-period framework.

APPENDIX

Proposition 1. If $U_{t+1}(C_t, w_{t+1} | \beta_{t+1})$ is monotone increasing and strictly concave (henceforth m.i.s.c.) in (C_t, w_{t+1}) , then $U_t(C_{t-1}, w_t | \beta_t)$ is m.i.s.c. in (C_{t-1}, w_t) .

Proof: The proof of the proposition relies primarily on straightforward applications of well-known properties of concave functions (cf. Iglehart). We first establish:

Lemma 1. If $U_{t+1}(C_t, w_{t+1} | \beta_{t+1})$ is m.i.s.c. in (C_t, w_{t+1}) , the expected utility function

$$\begin{aligned} & \int_{\beta_{t+1}} U_{t+1}(C_t, w_{t+1} | \beta_{t+1}) dF_{\beta_t}(\beta_{t+1}) \\ (10) \quad &= \int_{\beta_{t+1}} U_{t+1}(C_t, H_{\beta_t} R(\beta_{t+1})' | \beta_{t+1}) \\ & \quad \cdot dF_{\beta_t}(\beta_{t+1}) \end{aligned}$$

is strictly concave in (C_t, H_{β_t}) .

Proof: For any given value of β_{t+1} , and thus of $R = R(\beta_{t+1})$,

$$w_{t+1} = H_{\beta_t} R(\beta_{t+1})'$$

is a linear and thus concave (though not strictly concave) function of H_{β_t} . Since by assumption $U_{t+1}(C_t, w_{t+1} | \beta_{t+1})$ is m.i.s.c. in (C_t, w_{t+1}) , $U_{t+1}(C_t, H_{\beta_t} R(\beta_{t+1})' | \beta_{t+1})$ is strictly concave in (C_t, H_{β_t}) .¹⁸ Integrating over β_{t+1} in (10) preserves this concavity.

The remainder of the proof of Proposition 1 is then as follows. Let c_t^* , $H_{\beta_t}^*$ and \bar{c}_t^* , $\bar{H}_{\beta_t}^*$ be the optimal values of c_t and H_{β_t} in (4) for any two vectors (C_{t-1}, w_t) and $(\bar{C}_{t-1}, \bar{w}_t)$ that differ in at least one element. Let

$$\begin{aligned} C_{t-1} &= \alpha C_{t-1} + (1 - \alpha) \bar{C}_{t-1}, \\ w_t &= \alpha w_t + (1 - \alpha) \bar{w}_t, \\ c_t &= \alpha c_t^* + (1 - \alpha) \bar{c}_t^*, \\ \bar{H}_{\beta_t} &= \alpha H_{\beta_t}^* + (1 - \alpha) \bar{H}_{\beta_t}^*, \quad 0 < \alpha < 1. \end{aligned}$$

To establish the concavity of $U_t(C_{t-1}, w_t | \beta_t)$, we must show that

$$\begin{aligned} U_t(\bar{C}_{t-1}, \bar{w}_t | \beta_t) & > \alpha U_t(C_{t-1}, w_t | \beta_t) \\ & + (1 - \alpha) U_t(\bar{C}_{t-1}, \bar{w}_t | \beta_t). \end{aligned} \quad (11)$$

From Lemma 1, for $0 < \alpha < 1$,

$$\begin{aligned} & \int_{\beta_{t+1}} U_{t+1}(\bar{C}_{t-1}, \bar{c}_t, \bar{H}_{\beta_t} R(\beta_{t+1})' | \beta_{t+1}) dF_{\beta_t}(\beta_{t+1}) \\ & > \alpha \int_{\beta_{t+1}} U_{t+1}(C_{t-1}, c_t^*, H_{\beta_t}^* R(\beta_{t+1})' | \beta_{t+1}) \\ & \quad \cdot dF_{\beta_t}(\beta_{t+1}) \\ & + (1 - \alpha) \int_{\beta_{t+1}} U_{t+1}(C_{t-1}, \bar{c}_t^*, \bar{H}_{\beta_t}^* R(\beta_{t+1})' | \beta_{t+1}) \\ & \quad \cdot dF_{\beta_t}(\beta_{t+1}) \\ & = \alpha U_t(C_{t-1}, w_t | \beta_t) \\ & + (1 - \alpha) U_t(\bar{C}_{t-1}, \bar{w}_t | \beta_t). \end{aligned} \quad (12)$$

¹⁸ If $f(x_1, x_2, \dots, x_N) = f(X)$ is m.i.s.c. in X , and if $x_i = g_i(y_1, y_2, \dots, y_N) = g_i(Y)$, $i = 1, 2, \dots, N$, is concave (though not necessarily strictly concave) in Y , then $f(g_1(Y), g_2(Y), \dots, g_N(Y)) = f(G(Y))$ is strictly concave in Y . (See, e.g., H. G. Eggleston, p. 52.)

Since the consumption-investment decision implied by c_t , H_{β_t} is not necessarily optimal for the wealth level w_t ,

$$\begin{aligned} & U_t(C_{t-1}, w_t | \beta_t) \\ & \geq \int_{\beta_{t+1}} U_{t+1}(\bar{C}_{t-1}, \bar{c}_t, \\ & \quad \bar{H}_{\beta_t} R(\beta_{t+1})' | \beta_{t+1}) dF_{\beta_t}(\beta_{t+1}), \end{aligned}$$

which, with (12) implies (11).¹⁹

The monotonicity of $U_t(C_{t-1}, w_t | \beta_t)$ in (C_{t-1}, w_t) follows straightforwardly from the monotonicity of $U_{t+1}(C_t, w_{t+1} | \beta_{t+1})$ in C_t . Thus the proposition is established.

Finally, as noted earlier (fn. 10), when utility is written (as we have done throughout) as a function of consumption dollars, we are implicitly summarizing the consumption opportunities (in terms of goods and services and their anticipated prices) that will be available in each period. We shall now conclude the paper by showing how a von Neumann-Morgenstern "cardinal" utility function for consumption dollars can be derived from a cardinal utility function for consumption commodities.

Let $q(\beta_t) = (q_1, q_2, \dots, q_N(\beta_t))$ be the vector of quantities of $N(\beta_t)$ available commodities consumed during t in state β_t and let $p(\beta_t) = (p_1, p_2, \dots, p_N(\beta_t))$ be the corresponding price vector. In any period or state one of the available consumption commodities is always "dollar gifts and bequests" which has price \$1 per unit. At the horizon $\tau+1$, dollar gifts and bequests, denoted $w_{\tau+1}$, is the only available consumption good. Let

$$Q_\tau = (q(\beta_{1-\tau}), \dots, q(\beta_1), \dots, q(\beta_\tau))$$

be the vector representing lifetime consumption of commodities, and let $V(Q_\tau, w_{\tau+1} | \beta_{\tau+1})$ be the consumer's utility of lifetime con-

¹⁹ It is assumed that (c_t, H_{β_t}) is a feasible consumption-investment decision for the wealth level w_t , or equivalently, that the set of feasible values of (c_t, H_{β_t}) is convex. But this is a weak assumption that will be met, for example, when the constraints on c_t and H_{β_t} are equations such as $H_{\beta_t} c_t = w_t - c_t$ or linear inequalities such as $0 \leq c_t \leq w_t$ or $H \leq H \leq \bar{H}$, where \underline{H} and \bar{H} are vectors of lower and upper bounds on quantities invested in each asset.

sumption, given state β_{r+1} at $r+1$, and where $\beta_{1-k} \subset \dots \subset \beta_1 \subset \dots \subset \beta_{r+1}$. The utility function for *dollars* of consumption can then be defined as

$$(13) \quad U_{r+1}(C_{r+1} | \beta_{r+1}) = \max_{Q_r} V(Q_r, w_{r+1} | \beta_{r+1})$$

Subject to

$$C_{r+1} = (p(\beta_{1-k})q(\beta_{1-k})', \dots, p(\beta_r)q(\beta_r)', w_{r+1}) \\ = (c_{1-k}, \dots, c_r, c_{r+1}).$$

The role of β_{r+1} in U_{r+1} is twofold. First, psychological attitudes towards current and past consumption (or "tastes") may depend on the state of the world. Second, even if tastes for consumption *commodities* are not state dependent (so that β_{r+1} can be dropped from V), the utility of any stream of *dollar* consumption expenditures depends on the history of the set of available consumption commodities and their prices, both of which are subsumed in β_{r+1} .

A utility function $U_{r+1}(C_{r+1} | \beta_{r+1})$ which has the properties required by Proposition 1 can then be obtained from $V(Q_r, w_{r+1} | \beta_{r+1})$ as follows.

Proposition 2. If $V(Q_r, w_{r+1} | \beta_{r+1})$ is m.i.s.c. in (Q_r, w_{r+1}) , then $U_{r+1}(C_r, w_{r+1} | \beta_{r+1})$ is m.i.s.c. in (C_r, w_{r+1}) .

Proof: Let Q_r^* be the optimal value of Q_r in (13) for (C_r, w_{r+1}) and let \tilde{Q}_r^* be optimal for $(\tilde{C}_r, \tilde{w}_{r+1})$, where the vectors (C_r, w_{r+1}) and $(\tilde{C}_r, \tilde{w}_{r+1})$ differ in at least one element. For $0 < \alpha < 1$, let

$$(\hat{Q}_r, \hat{w}_{r+1}) \\ = \alpha(Q_r^*, w_{r+1}) + (1 - \alpha)(\tilde{Q}_r^*, \tilde{w}_{r+1}), \\ (\hat{C}_r, \hat{w}_{r+1}) \\ = \alpha(C_r, w_{r+1}) + (1 - \alpha)(\tilde{C}_r, \tilde{w}_{r+1}).$$

Then the strict concavity of V implies

$$V(\hat{Q}_r, \hat{w}_{r+1} | \beta_{r+1}) \\ > \alpha V(Q_r^*, w_{r+1} | \beta_{r+1}) \\ + (1 - \alpha) V(\tilde{Q}_r^*, \tilde{w}_{r+1} | \beta_{r+1}).$$

Or equivalently,

$$V(\hat{Q}_r, \hat{w}_{r+1} | \beta_{r+1}) \\ > \alpha U_{r+1}(C_r, w_{r+1} | \beta_{r+1}) \\ + (1 - \alpha) U_{r+1}(\tilde{C}_r, \tilde{w}_{r+1} | \beta_{r+1}).$$

Since \hat{Q}_r is a feasible but not necessarily an optimal allocation of \hat{C}_r , an optimal allocation must have utility at least as high as that implied by \hat{Q}_r , so that

$$U_{r+1}(\hat{C}_r, \hat{w}_{r+1} | \beta_{r+1}) \\ > \alpha U_{r+1}(C_r, w_{r+1} | \beta_{r+1}) \\ + (1 - \alpha) U_{r+1}(\tilde{C}_r, \tilde{w}_{r+1} | \beta_{r+1}),$$

and the concavity of U_{r+1} is established.

To establish the monotonicity of U_{r+1} in C_r , simply note that if the dollars available for consumption in any period are increased, consumption of at least one commodity can be increased without reducing consumption of any other commodity, so that utility must be increased. An optimal reallocation of consumption expenditures must do at least as well.

REFERENCES

- K. J. Arrow, "The Role of Securities in the Optimal Allocation of Risk-Bearing," *Rev. Econ. Stud.*, Apr. 1964, 31, 91-96.
- M. E. Blume, "The Assessment of Portfolio Performance: An Application of Portfolio Theory," unpublished doctoral dissertation, Grad. School of Business, Univ. Chicago 1968.
- P. A. Diamond, "The Role of A Stock Market in a General Equilibrium Model with Technological Uncertainty," *Amer. Econ. Rev.*, Sept. 1967, 57, 759-76.
- J. H. Drèze, "Fondements Logiques de la Probabilité Subjective et de l'Utilité," *La Decision*, Centre National de la Recherche Scientifique, Paris 1961, 73-87.
- and F. Modigliani, "Epargne et Consommation en Avenir Aleatoire," *Cahiers du Séminaire D'Econometrie*, 1966, 9, 7-33.
- H. G. Eggleston, *Convexity*. Cambridge, Mass. 1958.
- E. F. Fama, "The Behavior of Stock Market Prices," *J. Bus. Univ. Chicago*, Jan. 1965, 38, 34-105.

- , "Multiperiod Consumption-Investment Decisions," Report No. 6830, Center for Mathematical Studies in Business and Economics, Univ. Chicago, rev. May 1969.
- , "Portfolio Analysis in a Stable Paretian Market," *Manage. Sci.*, Jan. 1965, 12, 404-19.
- , (a) "Risk, Return, and Equilibrium," Report No. 6831, Center for Mathematical Studies in Business and Economics, Univ. Chicago, June 1968.
- , (b) "Risk, Return, and Equilibrium: Some Clarifying Comments," *J. Finance*, Mar. 1968, 23, 29-40.
- N. Hakansson, "Optimal Investment and Consumption Strategies for a Class of Utility Functions," Working Paper No. 101, Western Management Science Institute, Univ. California, Los Angeles, June 1966
- J. Hirshleifer, "Investment Decision Under Uncertainty: Applications of the State-Preference Approach," *Quart. J. Econ.*, May 1966, 80, 252-77.
- , "Investment Decision Under Uncertainty: Choice-Theoretic Approaches," *Quart. J. Econ.*, Nov. 1965, 79, 509-536.
- D. L. Iglehart, "Capital Accumulation and Production for the Firm: Optimal Dynamic Policy," *J. Manage. Sci.*, Nov. 1965, 12, 193-205.
- M. Jensen, "Risk, the Pricing of Capital Assets, and the Evaluation of Investment Portfolios," *J. Bus. Univ. Chicago*, Apr. 1969, 42, 167-247.
- J. Lintner, (a) "Security Prices, Risk, and Maximal Gains from Diversification," *J. Finance*, Dec. 1965, 20, 587-615.
- , (b) "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Rev. Econ. Statist.*, Feb. 1965, 47, 13-37.
- N. Liviatan, "Multiperiod Future Consumption as an Aggregate," *Amer. Econ. Rev.*, Sept. 1966, 56, 828-40.
- B. Mandelbrot, "The Variation of Certain Speculative Prices," *J. Bus. Univ. Chicago*, Oct., 1963, 36, 394-419.
- H. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*. New York 1959.
- J. Mossin, "Equilibrium in a Capital Asset Market," *Econometrica*, Oct. 1966, 34, 768-83.
- , "Optimal Multiperiod Portfolio Policies," *J. Bus. Univ. Chicago*, Apr. 1968, 41, 215-29.
- Edmund Phelps, "The Accumulation of Risky Capital: A Sequential Utility Analysis," *Econometrica*, Oct. 1962, 30, 729-43; reprinted in *Risk Aversion and Portfolio Choice*. D. Hester and J. Tobin, eds., New York 1967, 139-53.
- J. Pratt, "Risk Aversion in the Small and in the Large," *Econometrica*, Jan.-Apr. 1964, 32, 122-136.
- R. Roll, "Efficient Markets, Martingales, and the Market for U.S. Government Treasury Bills," unpublished doctoral dissertation, Grad. School of Business, Univ. Chicago 1968.
- L. Savage, *The Foundations of Statistics*. New York 1954.
- W. F. Sharpe, "Capital Assets Prices: A Theory of Market Equilibrium under Conditions of Risk," *J. Finance*, Sept. 1964, 19, 425-42.
- , "A Simplified Model for Portfolio Analysis," *J. Manage. Sci.*, Jan. 1963, 10, 277-93.
- J. Tobin, "Liquidity Preference as Behavior Towards Risk," *Rev. Econ. Stud.*, Feb. 1958, 25, 65-86.
- , "The Theory of Portfolio Selection." *The Theory of Interest Rates*. F. H. Hahn and F. P. R. Brechling, eds, London 1965. Ch. 1.
- J. von Neumann, and O. Morgenstern, *Theory of Games and Economic Behavior*, 3rd ed, Princeton 1953.

COMMUNICATIONS

The Life-Cycle Hypothesis: A Reinterpretation and Empirical Test

By MICHAEL LANDSBERGER*

The life-cycle hypothesis, generally referred to as the MBA theory after its formulators, Franco Modigliani, Richard Brumberg, and Albert Ando, has many points in common with the permanent-income hypothesis. The latter, however, has attracted considerably more attention in the literature, perhaps because Milton Friedman from the beginning suggested various tests to be performed in order to verify the empirical relevance of his theory. The life-cycle hypothesis was not accompanied by such suggestions for empirical tests and consequently most of the testing of the MBA model has been made by the authors themselves.

It is our opinion that the MBA model is somewhat ambiguous with respect to the hypothesis regarding the distribution over time of the marginal propensity to consume out of wealth. Since this hypothesis is probably the core of the MBA theory, the first task of this paper will be to attempt to clarify its logical place in the whole model. Then we will outline an approach to the problem of equivalent units which was recognized by Ando and Modigliani as an important issue in the MBA model. Our

suggestion, developed elsewhere, is oriented toward an empirical test which we perform and present in the second section of this paper.

The empirical test is based on data from two Saving Surveys conducted in Israel which cover the years 1957-58¹ and 1963-64. In these surveys we had information on windfall income (mainly restitution receipts)² which could be regarded as an increase in wealth equal to the windfall receipt. Therefore, the marginal propensity to consume out of windfall income, henceforth MPC^L , can be regarded as an estimate of the marginal propensity to consume out of wealth, MPC^w .

The organization of the paper is as follows: In Section I, we point out what in our opinion seems an ambiguity in the formulation of part of the MBA model. In Section II, we introduce briefly a definition of equivalent unit which seems necessary in presenting and testing the MBA model, and dwell on an empirical test of the MBA model as it concerns the idea of the life cycle in consumption. The results we arrive at support the conclusion of the MBA model.

I. The Distribution of MPC^L Over Time in the MBA Model

The MBA theory was first formulated by Brumberg and Modigliani in 1954, and

* This paper is based on a chapter of the author's Ph.D. thesis written under the supervision of N. Liviatan and submitted to the Hebrew University of Jerusalem. This paper, as other parts of the thesis, benefited from Professor Liviatan's guidance and the author would like to express his gratitude. F. Modigliani was very kind in reading an earlier draft of this paper and made several important comments, incorporated here, which were helpful in removing ambiguities and clarifying some points. I would also like to express my gratitude to the referee of this paper for his helpful comments. Responsibility for any remaining errors is, however, mine alone. The whole thesis was written as part of the author's work in the Research Department of the Bank of Israel.

¹ Two other surveys were conducted in 1958-59 and 1964-65, but they comprised a small number of families and therefore could not be useful for the test presented in this paper, which was based on the classification of families into subgroups.

² These were lump sums paid to Israeli families by the West German Government. The receipts were of order of magnitude of the current yearly income of the recipients.

later developed and amplified by Ando and Modigliani (1957, 1960, and 1964). Ando and Modigliani suggested the relaxation of some of the assumptions which appeared in the original Brumberg-Modigliani model, but these amendments do not affect the aspect of the theory that we intend to discuss in this paper.

The main points of the theory can be summarized with the following assumptions:

- (1) An individual seeks to maximize his utility function, which is homogeneous in present and future consumption, when the future extends over T periods³;

$$U = U(C_1, C_2 \dots C_T);$$

- (2) The interest rate is zero;
- (3) The individual does not inherit any wealth and does not intend to bequeath any.

Out of these MBA construct a consumption function of the form:

$$(4) \quad C_k = \gamma_k^t W_t, \quad k = t, t+1, \dots, T_t$$

Where:

C_k = consumption planned for period k

W_t = total wealth in period t

T_t = remaining life span at age t

γ_k = parameter

Since $r=0$ we may write the consumption function as

$$(5) \quad C_k = \gamma_k^t (Y_t + (N-t)Y_t^e + a_t)$$

where a_t represents nonhuman wealth in period t ; Y_t^e represents average income expected over the balance of the earning span $(N-t)$.

Finally, MBA add the assumption:

- (6) All the γ_k^t are equal, i.e., an individual plans to consume all his wealth at an even rate throughout the balance of his life. In this case, clearly, $\gamma_k^t = \gamma^t = 1/T_t$, and we may therefore write the consumption function as:

$$(7) \quad C_t = \frac{Y_t}{T_t} + \left(\frac{N-t}{T_t} \right) Y_t^e + \frac{a_t}{T_t}.$$

We may thus conclude that the marginal propensity to consume out of wealth equals $1/T_t$, which implies that this propensity rises with age; the older an individual is, the higher the value of $1/T_t$.⁴

We shall discuss this aspect of the consumption function here since it relates to windfall incomes and serves as a point of distinction between the MBA theory and others that regard wealth as the determinant of consumption (e.g. Friedman's theory). The distribution of the marginal propensity to consume out of wealth, by age, has a bearing on the idea of the savings cycle which is associated with MBA theory. Friedman, too, noted age as a variable influencing consumption, but the permanent-income theory does not lay any stress on this point. According to the MBA theory, the marginal propensity to consume out of wealth rises with age, therefore the marginal propensity to consume out of windfall income should also rise with age. Hence, data on windfall incomes are suitable for an efficient examination of the hypothesis concerning the distribution of the marginal propensity to consume out of wealth with age.

Before presenting our empirical test, we will discuss the way in which the hypothesis concerning the distribution of MPC^w or MPC^L over the life span derives from the MBA model. We deem it necessary to go into this matter because, in our opinion, the logical place of this hypothesis in the general MBA model is rather obscure.

On the whole, literature dealing with the consumption function tends to point at this hypothesis as the salient feature of MBA theory. Therefore, it is our opinion that if

³ In a later paper in 1963 (p. 57), Ando and Modigliani suggested the inclusion of planned bequeathals in the utility function, and claimed that this function is homogeneous in all consumptions and in the sum bequeathed.

⁴ We may assume that an individual's subjective life expectancy changes from time to time, so that the requirement $1/T_t < 1/T_{t+1}$ need not always be satisfied. Such an assumption can rest on the fact that there is a secular increase in life expectancy, which may affect an individual's subjective estimate. However if $T_t = f(t)$ and $dT/dt > 0$ it suffices to assume that $0 < dT_t/dt < 1$ which is a reasonable assumption.

this aspect of the MBA model is important, it should stem from assumptions that the proponents of MBA theory regard as essential to the model as a whole. In the opinion of MBA, the only such assumption is the one concerning the homogeneity of the utility function; but this assumption does not lead to any definite conclusions as regards the distribution of MPC^w over time, i.e., maximization of a homogeneous utility function with no additional specifications does not lead to any systematic variation in the marginal propensity to consume out of wealth (MPC^w) over time. The only assumption in the MBA theory that does, in our opinion, imply an increasing MPC^w over time is assumption (6); i.e., that an individual seeks to maintain a uniform level of consumption throughout his life. However, it should be noted that this is a sufficient but by no means a necessary assumption.* MBA, moreover, explicitly state (1963, p. 59), and (1954, pp. 395-6) that this assumption is not essential to their model, and is made merely to simplify the presentation. To sum up, the problem we are up against is that the hypothesis which we regard as important and central to MBA theory stems from an assumption that the writers themselves consider as unimportant, and one which they could do without. It is our opinion that one of two things must be done: either restrict the utility function further so that maximizing it will indeed cause the marginal propensity to consume out of wealth to rise with age; or accept assumption (6) as an important one in the model. The difference between the two alternatives is mainly formal, and not substantive.

In a letter to the author commenting on an earlier version of this paper, Modigliani suggested a third alternative, namely that he and his coauthors would be "... prepared to postulate that deviations of the preferred rate of consumption from a constant are not sufficiently large to invalidate the conclusion that MPC^w will tend to rise with age in the large." The condition for MPC^w to rise with age not only in the large but also in the small

is, according to Modigliani, "... that planned consumption should be a nondecreasing function of age..." This is, as Modigliani noted, a sufficient condition. "... The necessary and sufficient condition is that it should decrease no faster than implied by the inequality:

$$\frac{C_t}{C_s} > \frac{1}{W_t} \frac{dW_t}{dt} \dots "$$

where⁶

$$C_t \equiv \frac{dC_t}{dt}.$$

We accept this suggestion which, like ours, recognizes the deficiency in the original presentation of the model. We shall not discuss here which of the three modifications (assumption (6), restriction of the utility function, or the above mentioned one), should be introduced into the model in order to improve its consistency. In our opinion, the differences between the alternative suggestions are of minor importance.

We now turn to proving the contention that the above mentioned distribution of MPC^w over time does not necessarily follow from the maximization of a homogeneous utility function. We do this using a specific example. Assume that an individual has a homogeneous utility function

$$(8) \quad U = U(C_1, C_2, C_3) = C_1^\alpha C_2^\beta C_3^\gamma$$

of the order $(\alpha + \beta + \gamma)$. Only three periods were included, but in practice, C_2 can be regarded as representing the consumption of all the intermediate periods between the first and the last. We might add that since our purpose is to disprove a certain contention, we do not have to discuss the general case of T periods. From the maximization process of (8), under the constraint of (9):

* This inequality was obtained as a result of differentiation of $MPC^w(t) = C(t)/W(t)$ with respect to t and imposing

$$\frac{d\left(\frac{C(t)}{W(t)}\right)}{dt} > 0$$

* This reservation was raised by Modigliani.

$$(9) \quad C_1 + \frac{C_2}{1+r} + \frac{C_3}{(1+r)^2} = Y_1 + \frac{Y_2}{1+r} + \frac{Y_3}{(1+r)^2}$$

where the Y 's indicate present and future incomes, we get

$$(10) \quad \frac{C_1}{C_2} = \frac{\alpha}{\beta(1+r)}$$

$$(11) \quad \frac{C_2}{C_3} = \frac{\beta}{\gamma(1+r)}$$

Let us now assume that an individual maximizes his utility function one period later.⁷ Here we should note that the form of the utility function in the following period is not given automatically; even if we maintain that the function did not change, the very fact that an additional period has elapsed raises problems as to its new form. It would have been natural to write the new function as:

$$(12) \quad U = U(C_2, C_3) = C_2^\beta C_3^\gamma$$

since C_2 now indicates the consumption in the first period. But here the problem arises that Robert Strotz has already dealt with. Solutions of future consumption may differ according to the period in which they are arrived at, in spite of the fact that the utility function remains unchanged. Now Strotz has shown that solutions obtained in different periods will be identical under the conditions of a certain discount function. Such consistency seems a reasonable requirement of the utility function, and we therefore stipulate that the exponents of C_2 and C_3 in (12) must be such that the solutions obtained for C_2 and C_3 in the first period equal these obtained in the second period.⁸ Since we are discussing a homogeneous utility function in which only the ratios between consump-

tions are determined, we require $C_2/C_3 = C_2'/C_3'$ where C_2' and C_3' indicate consumption obtained by maximizing⁹ equation (12). Since $C_2/C_3 = \beta/\gamma(1+r)$ we require that $\beta'/\gamma' = \beta/\gamma$; i.e., the ratios between the exponents of C_2 and C_3 in the function

$$U = U(C_2, C_3) = C_2^{\beta'} C_3^{\gamma'}$$

should equal the ratios between the respective exponents in the function

$$U = C_1^\alpha C_2^\beta C_3^\gamma$$

Now let us assume that the individual is in the first period and receives a windfall of \$1 which he decides to consume during his life. The addition to consumption in each of the three periods is:

$$\Delta C_1, \Delta C_2, \text{ and } \Delta C_3.$$

Three conditions must be satisfied:

$$\left. \begin{aligned} \frac{\Delta C_1}{\Delta C_2} &= \frac{\alpha}{\beta(1+r)} \\ \frac{\Delta C_2}{\Delta C_3} &= \frac{\beta}{\gamma(1+r)} \end{aligned} \right\} \text{owing to homogeneity}$$

$$\Delta C_1 + \frac{\Delta C_2}{(1+r)} + \frac{\Delta C_3}{(1+r)^2} = 1,$$

owing to budget constraint.

Solving for ΔC_1 , we obtain

$$(13) \quad \Delta C_1 = \frac{1}{1 + \frac{\beta + \gamma}{\alpha}}$$

Alternatively, let us assume that the individual receives a windfall of \$1 in the second period and decides to consume all of it in his lifetime. In this case, the additions to present and future consumption must meet two conditions:

⁷ This may be regarded as a case in which the individual "forgot" the solutions for C_2 and C_3 obtained from (10) and (11), and solves them anew.

⁸ Although MBA did not mention this assumption explicitly, Modigliani agreed that it "... definitely underlies the MBA model...."

⁹ It is also possible to require $C_2 = C_2'$ and $C_3 = C_3'$ but for our purpose equality between the ratios suffices. We should remember that considering the ordinality of the utility function, the equality of ratios is no less strong a requirement.

$$\frac{\Delta C'_2}{\Delta C'_1} = \frac{\beta'}{\gamma'(1+r)}, \quad \text{owing to homogeneity,}$$

$$\Delta C'_2 + \frac{\Delta C'_1}{1+r} = 1,$$

owing to budget constraint.

Solving for $\Delta C'_2$ we get¹⁰

$$(14) \quad \Delta C'_2 = \frac{1}{1 + \left(\frac{\gamma'}{\beta'}\right)}$$

The question now is whether $\Delta C'_2 > \Delta C_1$ for all the possible values of $\alpha, \beta, \gamma, \beta', \gamma'$; the answer is no. To give a concrete example: let us assume $\alpha = 6\beta = 6\gamma$; then $\Delta C_1 = 3/4$, $\Delta C'_2 = 1/2$. This example is ample proof, we think, that a homogeneous utility function in itself does not cause the required regularity in the distribution of the marginal propensity to consume out of wealth over time.¹¹ In order to arrive at the conclusion concerning the distribution of the marginal propensity to consume out of wealth over time, we must either impose additional constraints on the homogeneous utility function or else, not use the utility function at all and instead use an assumption of the type¹² of equation (6). Note that we say "of the type of (6)" and not (6) as it appears in the MBA

¹⁰ It should be remembered that ΔC_1 can be regarded as an estimate of the marginal propensity to consume out of windfall income or wealth, in period t (MPC_t^w), whereas $\Delta C'_2$ represents an equivalent estimate for period $t+1$ (MPC_{t+1}^w) and according to the life-cycle hypothesis $MPC_{t+1}^w > MPC_t^w$.

¹¹ In this type of utility function the coefficients α, β and γ represent the elasticities of utility with respect to C_1, C_2 and C_3 , respectively. Thus, our concrete example reflects a very high elasticity of utility with respect to present consumption as compared with future consumption. Or, to put it on a more intuitive basis, our individual reveals a strong preference for present as compared with future consumption.

¹² An example of the kind of utility function mentioned here, that makes the marginal propensity to consume out of wealth rise with age, is:

$$U = U(C_1^\alpha C_2^\beta C_3^\gamma),$$

or when the function is symmetrical

$$U = U(C_1 C_2 C_3)^\alpha.$$

model, since the conclusion that MPC^w rises over time can be reached with a less limiting assumption, e.g., let us assume that the optimum consumption plan over time is

$$C, Cq, Cq^2, \dots, Cq^{T-1}$$

where $T-1$ is the planning span. In other words, an individual changes his consumption in each period relative to the previous one by q , where $q > 0$ and independent of wealth (assumption (6) is a specific case in which $q=1$).

Let us further assume that this individual receives \$1 windfall income and decides to divide it over the entire period without altering the relationship between consumptions in the various subperiods (this assumption is essential since we are still working under the overall assumption that the utility function is homogeneous). In this case, the additions to consumption will be $\Delta C, q\Delta C, q^2\Delta C, \dots, q^{T-1}\Delta C$, and owing to the budget constraint, there exists:

$$(15) \quad \Delta C \left(1 + \frac{q}{1+r} + \dots + \frac{q^{T-1}}{(1+r)^{T-1}} \right) = 1$$

$$\Delta C = \frac{1 - \frac{q}{1+r}}{1 - \left(\frac{q}{1+r} \right)^T}$$

Alternatively, let us assume that the individual receives \$1 windfall income one period later and decides to consume it as in the above example. Again, we get

$$(16) \quad \Delta C' = \frac{1 - \frac{q}{1+r}}{1 - \left(\frac{q}{1+r} \right)^{T-1}}$$

where $\Delta C'$ is the addition to consumption in the first period of the latter type.

The question now is: what is the relationship between ΔC and $\Delta C'$, a relationship that expresses the change with age in the marginal propensity to consume out of wealth. $\Delta C' > \Delta C$ means that the marginal propensity to consume out of wealth rises with

age, and vice versa; $\Delta C' < \Delta C$ means that it declines with age. Dividing ΔC by $\Delta C'$ we get:

$$(17) \quad \frac{\Delta C}{\Delta C'} = \frac{1 - \left(\frac{q}{1+r}\right)^{T-1}}{1 - \left(\frac{q}{1+r}\right)^T} = \frac{\sum_{t=0}^{T-2} \left(\frac{q}{1+r}\right)^t}{\sum_{t=0}^{T-1} \left(\frac{q}{1+r}\right)^t} < 1$$

since $q/(1+r) > 0$. This means that for every $q > 0$ the marginal propensity to consume out of wealth rises with age.¹³

II. An Empirical Test of the Variation of MPC^L with Age, and the Family Size Factor in the MBA Model

Any theory dealing with the optimum consumption path of an economic unit must give a precise definition of such a unit. In the absence of such a definition, it is assumed to be the family which is the basic unit in discussion aspects of consumption. The family itself, however, undergoes changes in the number of its members and its composition. The question is, therefore, whose consumption are we discussing? Modigliani and Brumberg did not deal with this subject in their first essay. Ando and Modigliani admit in their second essay (1957, a reply to Fisher's reservations) that the matter should be dealt with and is, indeed, essential to their theory. The authors repeated it again, (1960), but as far as the present author knows, no concrete suggestion was published.

It seems quite natural to us that the family's optimum consumption path has to be presented in terms of economic welfare the family enjoys. This economic welfare

derives from the family's consumption. We shall assume that a family is aware of its own economic welfare that stems from the level of consumption. We may maintain that there exists a basic unit, probably a married couple, that plans the family's desired distribution of consumption over the life span, while taking account of the distribution of the number of family members.

Now the question is: what should a family's consumption distribution be in order to maintain a desirable distribution of economic welfare? If, for example, a family intends to maintain a uniform level of welfare throughout the life span, what should be its consumption distribution over time considering expected changes in the number of its members. We assume that in each period it is possible to define a function E whose arguments are the number of family members and their composition so that there exists a utility function

$$U = U\left(\frac{C_1}{E_1}, \frac{C_2}{E_2}, \dots, \frac{C_T}{E_T}\right)$$

Where E_t is the value of E in period t , account being taken of the number of family members and their composition in each period. We shall refer to E_t as to the number of equivalent units.

Now, for the sake of simplicity and following the suggestion of MBA, let us assume that the family's aim is to maintain a uniform level of consumption for each equivalent unit over the planning span that extends over T periods.¹⁴

$$(18) \quad \frac{C_1}{E_1} = \frac{C_2}{E_2} = \dots = \frac{C_T}{E_T}$$

where:

C_t = total consumption of the family in period t

E_t = number of equivalent units in the family in period t .

We assume that (18) is to be maintained under the constraint:

¹³ This conclusion can be arrived at by maximizing a utility function, $U = C_1^\alpha C_2^\alpha \dots C_T^\alpha$. This formulation has an advantage in that it relates the demand function to the utility function. We chose not to use this utility function in our presentation because our purpose was to use a formulation similar to assumption (6).

¹⁴ This assumption of identical levels of consumption was made by MBA with no discussion, however, of changes in family size.

$$(19) \quad \sum_{i=0}^{T-1} C_{i+1}(1+r)^{-i} = W_1$$

where W_1 is the family's total wealth in the first period.

It can be proved (see Appendix 5 of Landsberger) that a unique consumption function can be derived from (18) and (19) and it is of the form

$$(20) \quad \frac{C_t}{E_t} = \frac{W_t}{\sum_{i=t-1}^{T-1} E_{i+1}(1+r)^{t-(i+1)}}$$

i.e., consumption per equivalent unit depends on total wealth divided by the discounted sum of equivalent units that will comprise the family to the end of the planning span.¹⁵ From (20) it follows that the MPC^L per equivalent unit, E , rises with age. It can also be proved (see Appendix 5 of Landsberger) that as long as there exists a nonnegative relationship between the number of family members and age, the consumption function formulated in (20) also implies that MPC^L per family rises with age. This is an important conclusion in this context because we have no data on equivalent units and shall be using data on families in our empirical test.

We found in our data a nonnegative relationship between age of family head and family size up to the age of 40-45; we therefore selected families of up to 45 years of age, and divided them into two groups: (a) families whose head is no older than 34; (b) families whose head is 35-44 years old. Division into a large number of groups was not possible because it would have resulted in too-small groups.

For each group we calculated the regression:¹⁶

¹⁵ Such discounting is necessary because the consumption of an equivalent unit that will be added in the j -th period is equal, from the viewpoint of its strain on the budget, to $1/(1+r)^j$ of present consumption.

¹⁶ The referee of this paper correctly remarked that insofar as these or other windfalls (such as restitution receipts) were anticipated, they might have been spent, partially at least, prior to receipt and this would tend to diminish the marginal propensity to consume out of such a windfall. On a priori grounds, this reservation can not be refuted. However, an empirical test we per-

$$(21) \quad C = \beta_{0g} + \beta_{1g}Y + \beta_{2g}P + \beta_{3g}O + \beta_{4g}A + V$$

where:

C = family consumption expenditure,
 Y = family current income,
 P = restitution receipts,
 O = other lump sums,
 A = age of head of family, and
 g = the index of the age group.

We distinguish between restitutions and other windfall incomes owing to their different effects on consumption and the relationship between the relative frequency of such receipts in various age groups. Since even after dividing the families into age groups, there still remained some variance in age, we introduced A into the regression.¹⁷ Actually, A should be introduced into the equation together with an expression of its interaction with the income and wealth variables (see Ando and Modigliani 1957, p. 113). This is the reason why it is convenient to use a classification into age groups when dealing with the effect of age. In our case, however, most of the variance in age is taken care of by that classification so that the residual variance which the variable A is supposed to measure is small. Equation (21) was calculated from the 1957-58 and 1963-64 saving surveys within each age group, 18-34, and 35-44; from it we obtain two estimates of the family's marginal propensity to consume out of wealth in each age group: the

formed revealed, according to our opinion, that neither restitutions nor other windfall receipts employed in this paper, affect consumption until they are received. Out of two saving surveys, we had information on identical families for two consecutive years. Using these data we checked whether consumption in period t was affected by windfall receipts in period $t+1$. The results obtained did not reveal any such effect. Therefore, we think that the whole question of anticipation can be disregarded at the practical level. The above mentioned test was performed by a regression analysis.

¹⁷ The entire discussion on the effect of age involves the problem of how to make a quantitative estimate of a family's age, age being an attribute of an individual. There does not seem to be a better solution than taking the age of the head of the family. There might be superior theoretical methods, but for lack of a practical way in which to apply them, we have not mentioned them at all.

TABLE 1—MARGINAL PROPENSITIES TO CONSUME OUT OF WINDFALL INCOME, BY AGE GROUPS

Age Group of family head	<i>MCP^a</i> Out of Personal Restitutions				<i>MCP^a</i> Out of Other Windfall Income ¹			
	Date of Savings Survey				Date of Savings Survey			
	1963-64		1957-58		1963-64		1957-58	
	Total consumption	Non-durable consumption	Total consumption	Non-durable consumption	Total consumption	Non-durable consumption	Total consumption	Non-durable consumption
18-34	0.2066 (±0.0503)	0.0597 (±0.0313)	0.0736 (±0.0431)	0.0142 (±0.0399)	0.2357 (±0.0500)	0.0727 (±0.0311)	0.1472 (±0.0336)	0.0305 (±0.0311)
35-44	0.1416 (±0.0379)	0.0620 (±0.0285)	0.2237 (±0.0411)	0.1500 (±0.0487)	0.1133 (±0.0762)	0.0394 (±0.0574)	0.2919 (±0.0565)	0.2372 (±0.0532)
Student's $ t $	1.03	0.05	2.52	2.16	1.34	0.57	2.20	3.34

Sources: The results presented in this table are based on data from the 1957/58 and 1963/64 Saving Surveys conducted in Israel. For further information see Central Bureau of Statistics and Bank of Israel research department.

¹ These were mainly comprised of windfalls such as inheritance, lotteries and prizes, money gifts, etc. All of them represent an increase in the wealth of the recipient equal to the windfall receipt. The referee of this paper pointed out that the distribution of *MPC^a* over age groups could be caused by a different composition of various lump sums in various age groups. In order to check this point we estimated the *MPC^a* using only one kind of windfall (which comprised about 70 percent out of total nonrestitution windfalls). The results disproved this reservation.

coefficients of *P* and *O* (the 1957-58 survey does not give data for *A*). According to the conclusions drawn in the beginning of this section we may expect:

$$\beta_{1,0} < \beta_{1,0+1} \quad \text{and} \quad \beta_{2,0} < \beta_{2,0+1},$$

i.e., we require that each of the coefficients of

P and *O* should be higher when estimated from data on families belonging to an older age group.

The results presented in Table 1 show that in five out of eight cases, the *MPC^a* rises with age. Looked at this way, the results are hardly encouraging. However, a

TABLE 2—MARGINAL PROPENSITIES TO CONSUME OUT OF WINDFALL INCOME, BY AGE GROUPS

Age Group	<i>MCP^a</i> Out of Personal Restitutions				<i>MCP^a</i> Out of Other Windfall Income			
	Date of Savings Survey				Date of Savings Survey			
	1963-64		1957-58		1963-64		1957-58	
	Total consumption	Non-durable consumption	Total consumption	Non-durable consumption	Total consumption	Non-durable consumption	Total consumption	Non-durable consumption
45-54	0.1268 (±0.0311)	0.0768 (±0.0268)	0.2608 (±0.0509)	0.1910 (±0.0496)	0.0498 (±0.1163)	0.0601 (±0.1003)	0.0569 (±0.0528)	0.0633 (±0.0513)
55+	0.1597 (±0.0146)	0.1098 (±0.0127)	0.1678 (±0.0191)	0.1274 (±0.0187)	0.1845 (±0.0597)	0.1893 (±0.0504)	0.2956 (±0.0777)	0.3089 (±0.0790)
Student's $ t $	1.00	1.11	1.71	1.22	1.03	1.15	2.54	2.61

Sources: See Table 1.

simple statistical test revealed that the differences in the MPC^L estimates between age groups are statistically significant only in four cases, and in all of them the MPC^L rises with age. Therefore we conclude that the results accord with what would be expected from the MBA theory.

Equation (21) was computed also for the remaining families, which were divided into two age groups, 45–54 and over 55 years. The significance of a comparison of MPC^L estimated from these two groups is restricted, however, since the average family size in the latter group is lower than in the former. Therefore, the family size effect acts in an opposite direction as compared with the effect expected from the life-cycle hypothesis. Thus this test can only confirm the MBA model. Unfavorable results can always be accounted for by the family size effect.

Looking at the results presented in Table 2, it appears that in spite of the family size effect, the MPC^L rises with age in six out of eight cases. However, the differences are significant only in two cases, in both of which the MPC^L rises with age, confirming once again the life-cycle hypothesis.

REFERENCES

A. Ando and F. Modigliani, "Tests of the Life Cycle Hypothesis of Saving," *Bull. Oxford Univ. Inst. Econ. Statist.* May 1957, 19, 99–123.

——— and ———, "The Permanent Income and the Life Cycle Hypothesis of Saving Behaviour, Comparison and Tests," published in *Proceedings of the Conference on Consumption and Savings*, Vol. II, Philadelphia, 1960.

——— and ———, "The 'Life Cycle' Hy-

pothesis of Saving: Aggregate Implications and Tests," *Amer. Econ. Rev.*, Mar. 1963, 53, 55–84.

——— and ———, "The 'Life Cycle' Hypothesis of Saving: A Correction," *Amer. Econ. Rev.*, Mar. 1964, 54, 111–13.

R. J. Ball and P. Drake, "The Relationship between Aggregate Consumption and Wealth," *Int. Econ. Rev.*, Jan. 1964, 5, 63–81.

R. Brumberg and F. Modigliani, "Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data," in *Post Keynesian Economics*, K. K. Kurihara, ed., New Brunswick 1954.

M. K. Evans, "The Importance of Wealth in the Consumption Function," *J. Polit. Econ.*, Aug. 1967, 75, 335–51.

M. J. Farrell, "The New Theories of the Consumption Function," *Econ. J.*, Dec. 1959, 69, 678–95.

M. R. Fisher, "Exploration in Savings Behavior," *Bull. Oxford Univ. Inst. Econ. Statist.*, Aug. 1956, 18, 201–77.

M. Friedman, *A Theory of the Consumption Function*, Princeton, 1957.

———, "Windfalls, the 'Horizon' and Related Concepts in the Permanent Income Hypothesis," in C. Christ, ed., *Measurement in Economics*, Stanford 1963.

M. Landsberger, *The Consumption Function and Windfall Receipts*, Bank of Israel, Research Department, Jerusalem, forthcoming.

A. Spiro, "Wealth and Consumption Function," *J. Polit. Econ.*, Aug. 1962, 70, 339–54.

R. Strotz, "Myopia and Inconsistency in Dynamic Utility Maximisation," *Rev. Econ. Stud.*, 23, 1955–6, No. 3, 165–80.

Central Bureau of Statistics and Bank of Israel, Research Department, *Saving Survey 1963–64*, Jerusalem 1967.

Labor Migration and Urban Unemployment: Comment

By PAUL ZAREMBKA*

The recent article by Michael P. Todaro on urban unemployment in less-developed countries is an important advance in the study of this pressing problem. The notion that rural workers take into account their probability of unemployment when moving to an urban center is an essential concept. Unfortunately, Todaro's paper contains a mathematical error which leads to a substantial overestimation of urban unemployment. Furthermore, the article also seems to make an error of specification which, more fortunately, is not crucial.

Todaro begins by assuming that "...the percentage change in the urban labor force as a result of migration during any period [$\dot{S}/S(t)$] is governed by the differential between the discounted streams of expected urban [$V_u(t)$] and rural [$V_R(t)$] real income" (p. 141), which he formulates as¹

$$(1) \quad \frac{S}{S}(t) = F \left[\frac{V_u(t) - V_R(t)}{V_R(t)} \right].$$

However, this equation seems to be an implausible specification since it implies that the level of migration does not depend upon rural population. In fact, the income differential operates on rural workers, so the equation should express migratory flow in terms of these workers:

$$(1a) \quad \frac{A}{A}(t) = F \left[\frac{V_u(t) - V_R(t)}{V_R(t)} \right]$$

where \dot{A} represents migration from agriculture while A represents agricultural labor

force. For example, equation (1a) may state that one in ten workers leaves agriculture when the expected income differential is 100 percent.

Now, since $\dot{A} = \dot{S}$ (emigration equals immigration), the equation for the growth in urban labor force due to migration becomes

$$(1b) \quad \frac{S}{S}(t) = \frac{A(t)}{S(t)} F \left[\frac{V_u(t) - V_R(t)}{V_R(t)} \right]$$

which differs from Todaro's by the coefficient $A(t)/S(t)$, the ratio of rural to urban labor force at time t . Of course, the function F has a somewhat different interpretation in this formulation. This respecification, as it turns out, only has substantial effects on the model after the urban sector has grown very considerably.

Todaro next specifies rural expected real income as

$$(2) \quad V_R(0) = \int_{t=0}^n Y_R(t) e^{-rt} dt$$

where $Y(t)$ is real income in period t , r is the rate of discount, and n is the time horizon. Urban expected real income is

$$(3) \quad V_u(0) = \int_{t=0}^n p(t) Y_u(t) e^{-rt} dt - C(0)$$

where $p(t)$ is the probability of finding an urban job in period t and $C(0)$ is moving costs. Then, after formulating the probability of finding an urban job, Todaro turns to a version of the model which has a one-period time horizon so that

$$(2a) \quad V_R(t) = Y_R(t)$$

and, since $p(t) = \Pi(t)$ in his one-period horizon case,

$$(3a) \quad V_u(t) = \Pi(t) Y_u(t),$$

assuming $C(0) = 0$.

Todaro makes the mathematical error in

* Assistant professor of economics, University of California, Berkeley. The author's research is currently being financed by the Ford Foundation and the National Science Foundation under the Project for the Explanation and Optimization of Economic Growth, Institute of International Studies, University of California, Berkeley.

¹ Equation numbers are the same as Todaro's. If they are followed by a letter, they originate with this paper.

his expression (8) which gives an equation for growth of aggregate labor supply in the urban area as

$$(8) \quad \frac{\dot{S}}{S}(t) = \beta + \Pi(t)F\left[\frac{Y_u(t) - Y_R(t)}{Y_R(t)}\right]$$

where \dot{S}/S now includes the natural rate of increase in urban population β . This equation should have been written

$$(8a) \quad \frac{\dot{S}}{S}(t) = \beta + F\left[\frac{\Pi(t)Y_u(t) - Y_R(t)}{Y_R(t)}\right]$$

as a result of substituting (3a) and (2a) into (1). This simple oversight is shown below to have drastic effects on the solution.

Making this correction and at the same time correcting the specification of migration (see equation (1b)), equation (8) becomes

$$(8b) \quad \frac{S}{S}(t) = \beta + \frac{A(t)}{S(t)}F\left[\frac{\Pi(t)Y_u(t) - Y_R(t)}{Y_R(t)}\right].$$

With Todaro, we solve for the equilibrium rate of employment in the urban sector E^* where

$$(10) \quad E(t) = \frac{N(t)}{S(t)}$$

and $N(t)$ is urban employment. Therefore, we solve for

$$(11) \quad \frac{E}{E}(t) = \frac{\dot{N}}{N}(t) - \frac{\dot{S}}{S}(t) = 0.$$

Now, Todaro posits the rate of urban employment growth as a constant γ and posits

$$(6) \quad \pi(t) = \frac{\gamma N(t)}{S(t) - N(t)} = \frac{\gamma E(t)}{1 - E(t)}.$$

¹ This expression does *not* imply that a 100 percent probability of finding a job in the initial period will lead to a 100 percent employment rate. In fact, the employment rate in this case is $1/(1+\gamma)$ and departs from one due to short-term unemployment (less than one time period). Note also that γ should be defined as the rate of growth in employment opportunities so that it includes demand for new employees due to attrition of old employees as well as increases in employment levels.

Using this and equation (8b), equation (11) becomes

$$\gamma - \beta - \frac{A(t)}{S(t)}F\left[\frac{\gamma E(t)}{1 - E(t)} \frac{Y_u(t)}{Y_R(t)} - 1\right] = 0.$$

If we suppose $F(X) = \eta X$ where η is a constant, then this last equation can be solved to give

$$(14a) \quad E^* = \frac{\gamma - \beta + \eta \frac{A(t)}{S(t)}}{\gamma - \beta + [1 + \alpha\gamma]\eta \frac{A(t)}{S(t)}}$$

where α is the ratio of urban to rural real income (not quite the same definition as Todaro's) with $\alpha > 1$. This equation has a close approximation: Since (i) the ratio of rural to urban population in less-developed countries is likely to be in the order of three and (ii) η might be about 0.10 (10 percent of the rural population would emigrate if the *expected* urban earnings were 100 percent higher) or more, then $\gamma - \beta$ is small relative to the other terms. Therefore, the solution for E^* is closely approximated by²

$$(14b) \quad E^* = \frac{1}{1 + \alpha\gamma}.$$

Todaro's solution for E^* in our definition of α is

$$(14) \quad E^* = \frac{\gamma - \beta}{\gamma F(\alpha - 1) + \gamma - \beta}.$$

As an example he takes $\gamma = 0.04$, $\beta = 0.02$, $F(\alpha - 1) = \alpha - 1$, and $\alpha = 2.00$ which gives $E^* = 0.33$; only one-third of the urban labor force is employed. However, our solution gives $E^* = 0.93$. Thus, a 67 percent unemployment rate is reduced to merely 7 percent.

Todaro's error has the effect of drastically raising his unemployment rate from the correct value. In its correct form $E^* = 1/(1 + \alpha\gamma)$, we find the unemployment rate much less and the probability of finding a job in

² Actually, if $\gamma - \beta > 0$, equilibrium employment is slightly greater than this approximation.

the urban area in the first time period much greater.⁴

Our solution for E^* indicates first that the actual form of the function F has little effect on the equilibrium employment rate. This conclusion is due to the approach of the expected urban income to the rural income so that the argument for F approaches zero.

Second, the solution for E^* shows that an improvement in employment opportunities in the urban sector, say through output expansion, will *increase* the unemployment rate through the resultant initial increase in the probability of finding employment and thus in migration. This conclusion is opposite the implication of Todaro's solution (p. 145). In fact, it is the correct solution which fits his empirical evidence for Kenya (p. 140).

Third, as Todaro comments (p. 145), a decrease in the real earnings differential between the sectors lowers the unemployment rate, but not as sharply as in Todaro's solution.

Finally, and perhaps most importantly, in order to keep the unemployment rate constant, the income ratio α and the rate of em-

ployment opportunities γ are *inversely* related, not directly, as Todaro claims to show (pp. 145-6). For example, to keep the employment rate constant, an increase in urban employment opportunities through output expansion must be accompanied by a *decline* in the ratio of urban to rural wage rates—a very difficult task since industrialization is often accompanied by a rise in urban wages. Thus, it seems that urban expansion is also of necessity going to be accompanied by more unemployment, at least until the urban sector becomes much more important than the rural (see equation (14a)).

A very interesting implication of the solution for $E^* = 1/(1+\alpha\gamma)$ is that, if we have data on the rate of urban employment growth, the rate of employment can be predicted from the real income differential or the income differential can be predicted from the employment rate. However, equation (14a) demonstrates that as the rural sector becomes small the unemployment rate approaches zero.

REFERENCE

- M. P. Todaro, "A Model of Labor Migration and Urban Unemployment in Less Developed Countries," *Amer. Econ. Rev.*, Mar. 1969, 59, 138-48.

⁴ The probability of finding an urban job is equal to $1/\alpha$ (see equation (6)) so that the expected income differential equals zero; it is much less in Todaro's, $(\gamma-\beta)/(\alpha-1)$ in his case when $F(\alpha-1)=\alpha-1$.

Labor Migration and Urban Unemployment: Reply

By MICHAEL P. TODARO

I would like to thank Paul Zarembka for pointing out what is apparently a very careless mathematical error in my paper. In actual fact, however, my carelessness was not so much a failure to make a correct algebraic substitution as a failure to explain briefly why I changed the form of my analytical equilibrium model equation (8) from that *directly* implied by the underlying behavioural model—i.e. why I did not use Zarembka's equation (8a). The model set forth in my paper represented an attempt to provide a concise and mathematically rigorous formulation of a phenomenon which was described verbally, in considerably more detail, in an earlier paper published in the *Yale Economic Essays*. Unfortunately, in my desire to be concise I carelessly forgot to point out in the sentence before equation (8) that for analytic as well as policy purposes I was separating the employment probability variable, $\pi(t)$, from the percentage urban-rural wage differential variable, $\alpha(t)$, so that each could be treated independently—i.e., the sentence should have read, "Next we specify an aggregate labor supply equation which is a simplified version of equation (1) in the sense that only a one-period time horizon is assumed and the probability variable (π) for analytical and policy purposes is treated separately from the wage differential variable." Now, having expressed my *mea culpa* for this carelessness, let me turn to Zarembka's correction and show why he also has been very careless in greatly exaggerating the quantitative significance of my apparent mathematical error. I shall then show why I feel that my equation (8) is a much better way of formulating the labor supply function than is Zarembka's (8a).

Briefly, Zarembka shows that if I had made the proper substitution in equation (8) on the basis of my earlier equations (2) and (3), then my equilibrium employment rate

could be "closely approximated" by

$$(14b) \quad E^* = \frac{1}{1 + \alpha\gamma}$$

rather than my derived equation

$$(14) \quad E^* = \frac{\gamma - \beta}{\gamma F(\alpha) + \gamma\beta}.$$

Zarembka then uses the same hypothetical parametric numbers as in my original paper to show that with his (14b) $E^* = 0.93$ whereas my equation (14) yields $E^* = 0.33$. "Thus, a 67 percent unemployment rate is reduced to merely 7 percent" and "Todaro's error has the effect of drastically raising his unemployment rate from the correct values." Unfortunately, what Zarembka fails to point out is that my result depends very largely on the nature of $F(\alpha)$, the migration response function. In my paper, I very arbitrarily assumed for purposes of explication and, "for simplicity," that $F(\alpha) = \alpha$ so that in my arithmetic example $F(\alpha) = 1.0$. Now, Zarembka assumes (and, I think, more realistically) that $F(\alpha) = \eta\alpha$ where η is a constant which "might be about 0.10." Therefore, if I were to substitute Zarembka's more realistic $F(\alpha) = 0.10\alpha$ for my simplistic $F(\alpha) = \alpha$ in equation (14), my new equilibrium E^* would equal 0.833 which is not "drastically" different from Zarembka's 0.93. It all depends on the specification of the function F , which leads me to my next point.

Zarembka argues on the basis of his "correct" equilibrium formula that "... the function F has little effect on the equilibrium employment rate" because of "... the approach of the expected urban income to rural income so that the argument for F approaches zero." While it is true that expected urban income approaches rural income in equilibrium, note that in our dynamic equilibrium there is still rural-urban

migration in an amount $\gamma - \beta$. Thus, although the approach to equilibrium implies a narrowing of the urban-rural expected income gap, such a gap is never completely eliminated. Moreover, even if one did allow the argument of P to equal zero, it does not follow on any a priori grounds that P will then also equal zero. Obviously, the precise nature of the migration response function is an important empirical question (and, let me point out, one which we are presently trying to estimate in Kenya on the basis of a very large sample survey of migrants). I certainly think that it should be in the equilibrium solution and this is one of the reasons why I prefer equation (14) over (14b).

Let me turn now to what I believe is a much more fundamental question raised by Zarembka's equation (14b). Once we have shown that Zarembka's "exaggeration" argument is incorrect, the basic difference that remains between Zarembka's (14b) and my equation (14) is that the former implies, *ceteris paribus*, that the higher the rate of urban job creation, the higher will be the rate of urban unemployment, whereas my formula states the opposite. The crux of my original argument was that it is not enough to attack the problem of urban unemployment merely through accelerated industrial growth without a concomitant drive to hold the line on the urban-rural wage differential (note, I did not say that one *must* reduce the differential as is implied by equation (14b)). I was not then, and am still not now prepared to go to the extreme and, I think, empirically incorrect statement that the urban unemployment rate will rise whenever the rate of employment creation

expands. Even a cursory glance at the data for, say, Seoul, Tokyo, Taipei, and Mexico City as opposed to those of, say, Calcutta, Lima, and Nairobi would support my interpretation.

Finally, let me point out that my model was really a one-sector partial equilibrium model which completely disregarded the size and nature of the rural sector. However, in an extension of this paper which takes into account the size of the rural and urban populations as well as rural production possibilities in the form of a two sector model, it is demonstrated (see Harris and Todaro p. 21) that the equilibrium urban unemployment rate *must necessarily decline* when the rate of employment creation is increased even though the *absolute number* of urban unemployed can increase. In my reference to the 1964 "Tripartite Agreement" in Kenya which Zarembka seems to think provides direct support for his conclusions, I quoted Professor Harbison as claiming that "the *volume* of unemployment . . . was probably increased rather than decreased." Neither he nor I ever claimed that the *rate* of urban unemployment was ever increased, a necessary conclusion of the Zarembka formula.

REFERENCES

- J. R. Harris and M. P. Todaro, "Migration, Unemployment and Development: A Two-Sector Analysis," *Amer. Econ. Rev.*, Mar. 1970, 60, 126-42.
- M. P. Todaro, "An Analysis of Industrialization, Employment, and Unemployment in Less Developed Countries," *Yale Econ. Essays*, No. 2, fall 1968, 8, 329-402.

The 1967 Peruvian Exchange Crisis: A Note

By LAWRENCE B. MORSE*

During the third quarter of 1967 the Peruvian sol underwent a major devaluation, the third in the past fifteen years. With the economy racked by inflation, the President's coalition rapidly deteriorating amid faint whispers of a military coup, and Central Reserve Bank's reserves becoming critically low, speculators had a heyday in the exchange markets. Recognizing that dollar reserves were nearly exhausted and that putting off the fateful day only helped prolong the "fundamental disequilibrium," on September 1 the Bank finally withdrew from the exchange market, abandoning the 26.82 soles/dollar peg which had been in effect since January 1961. The exchange market, freed of Central Reserve Bank intervention, depreciated precipitously: by the end of September the rate had fallen 40 percent. Then on October 5 by Supreme Decree, the President reestablished the certificate exchange system and the following day the Bank announced that when it opened for business on Monday, October 9, it would stand ready to buy and sell dollar certificates at 38.70 soles. The 44.3 percent depreciation appeared to have quelled the exchange crisis.

The purpose of this note is to examine the exchange policies which came out of this 1967 crisis. Essentially three policy decisions were made in October: to reestablish the dollar certificate market; to peg the certificate rate; and to peg the rate at 38.70 soles to the dollar. Though the confines of this comment do not allow an analysis of the second, and clearly, most basic action, there is evidence to suggest that the decision to peg the rate was an unfortunate

one.¹ As to whether 38.70 was an appropriate rate, this problem must await the accumulation of further evidence. We shall address ourselves to the first question, i.e., should the certificate market have been reestablished? It is our contention that it should not have.

Before presenting the arguments to be made against the reestablishment of the certificate system, a brief description of the two exchange markets is in order.

With the establishment of the certificate market in September 1948, exporters were required to obtain export licenses. The requirement was to insure that exporters surrendered their exchange to the Central Reserve Bank in return for which they were issued exchange certificates. Initially, certificates were denominated in dollars, pounds sterling, French francs and Argentine pesos. Exporters' exchange proceeds in other currencies were either allowed to be sold in the draft market or were issued dollar certificates at the appropriate cross rate. The volume of trade in these other currencies was, however, a very small portion of total exchange transactions. These certificates were, for a specified length of time, freely negotiable in the open market and were purchased by importers and those making certain types of capital payments abroad. Policing the use of exchange certificates was left to the commercial banks which required proof of the use to which the exchange was to be put. Repayment and servicing of foreign loans, repatriation of profits, payment of shipping and insurance charges, etc. were considered legitimate capital movements and could, on permission, be financed through the certificate market. Importers then surrendered the certificate, receiving

* Department of economics, Tougaloo College. The author is indebted to the Economic Development Center of the University of Minnesota for having financed the research and writing of this study, and to the Economic Growth Center of Yale University for the excellent research facilities which made this paper possible.

¹ See L. Morse. On the basis of an examination of the exchange policies from 1948 through 1965, it is the thesis of the study that various arguments can be made for Peru's operating under a flexible rather than a fixed exchange rate system.

the exchange needed for their international transactions. Thus the certificate market was fed by exporters' exchange proceeds and certain capital inflows and was used to finance imports and certain acceptable capital movements. From 1950 until it was temporarily suspended in May 1960, the certificate market generally accounted for four-fifths of total exchange transactions.

The draft market was—and remains—a free exchange market in the sense that no restrictions were placed on the use of exchange secured in that market. The market was fed by invisibles and certain capital inflows, and when the certificate market was suspended it carried all exchange transactions.

There are three arguments to be made against the reestablishment of the certificate system. The first, and perhaps strongest, is the simple observation that any form of exchange control results in resource misallocation. Second, evidence suggests that the certificate market never fully performed its intended function of stabilizing the primary exchange rate. Finally, the conditions which pertained at the inception of the certificate system in 1948 were not present in 1967. This last argument will serve as our starting point as it should allow some insights into the rationale of the certificate market.

In keeping with the practices of the time, from 1945 until its forceful removal from office in the latter part of 1948, the Bustamante administration saddled the Peruvian economy with ever stricter exchange and import controls. When the certificate system was introduced in September 1948, it represented a form of exchange control liberalization in spite of the fact that it initially meant the creation of a third exchange market. In a series of three steps, by November 1949, the certificate market had phased out the grossly overvalued and highly restricted official exchange market. Other import controls were also eased such that by January 1951, the list of prohibited imports had been totally abolished.

Although the exchange liberalization did not go all the way to a unification of the draft and certificate markets, nonetheless

the trend was clearly away from the immediate post-World War II controls, and the replacement of the official market with the certificate market was one of the most important aspects of that trend.

The reestablishment of the certificate market in 1967 was different from the 1948 inception in two important respects. First, it came as a form of exchange control rather than exchange liberalization. Second, it ran counter to the internationally fashionable practice of trade and exchange liberalization. Thus while the creation of the certificate market had helped Peru keep in step with international trends, the reestablishment of the system was a distinct and unfortunate regression.

The initial intent of the certificate market was “. . . *the avoidance of capital flight*, while permitting other proceeds from exports and other inflows, converted into certificates, to be used exclusively for the payment of imports and other normal current account services,” (emphasis added, see Central Reserve Bank of Peru). In practice, however, most capital account transactions could be transacted in the certificate exchange market. Other, and presumably “undesirable,” capital movements were financed through the draft market. The free movement of capital was deemed important in attracting foreign capital to Peru, yet at the same time the authorities did not want capital flight to disrupt the normal course of the exchange market which was to service the country's legitimate transactions; the assumption being that capital movement in the draft market would in no way distort the rates in the certificate market. In principle, the rates were to be independent of one another.² Thus during times of crisis when capital flight might become sufficiently large to distort the normal market forces under a unified exchange system, the draft market was to serve as a safety valve for the certificate market.

The rates were not totally independent of one another insofar as draft exchange could be used for any purpose while certificate exchange was restricted as to use, hence the draft rate was necessarily greater than, or equal to, the certificate rate.

TABLE 1—MEAN ANNUAL SPREADS BETWEEN DRAFT AND CERTIFICATE EXCHANGE RATES, COMPUTED FROM MONTHLY DATA*

Year	Spread (Percent)	Year	Spread (Percent)
1950	3.58 ^b	1955	1.15
1951	0.88	1956	1.33
1952	1.03	1957	0.54
1953	0.92	1958	0.97
1954	2.19	1959	0.79

* Source: *Memoria y Estadística*, 1950–59, Superintendent of Banks, Lima, Peru.

^b The spread was unusually large in 1950 due to the fact that the system was still readjusting to the November 1949 abandonment of the official exchange market and secondly, net long term-private foreign direct investments were negative that year.

In actuality, however, the spread between the two rates was never very large. The overall annual mean for 1950–59 was 1.34 percent. There are several factors which account for this. The most important was the fact that the certificate market was never a very restricted market, e.g., from 1952 on it carried better than 80 percent of all exchange transactions. Most capital transactions were granted access to the certificate market. Hence under normal circumstances there was relatively little need for the draft market as an escape valve. Further, there were two forms of arbitrage between the two markets.³ First, foreign investors served as the arbitrageurs as they sold their foreign exchange in the draft market and then used their sol proceeds to finance their capital goods imports through the certificate market. Such operations augmented the supply in the dear market and increased the demand in the cheaper market. One very simple measure of this arbitrage suggests that, *ceteris paribus*, a one million dollar increase in net long-term private foreign direct investment would be associated with a decrease in the spread of

³ The use of foreign exchange secured through prepayment for exports could have also acted as a form of arbitrage between the two markets. However, such operations would not necessarily tend to equalize the spot draft and spot certificate rates, except when the futures draft rate was expected to equal the futures certificate rate.

one-tenth of a percentage point.⁴

The second form of arbitrage is a rather traditional one between a restricted and an unrestricted market, namely cheating. That is, as the difference became larger it increased the incentive to exporters to under-invoice their exports and to importers to over-invoice their imports. Both devices freed exchange from the certificate market and allowed its clandestine sale in the premium draft market. However, since the differential involved was never very large, it is probable that this form of arbitrage was relatively unimportant given the risks involved.⁵

In view of these circumstances it is not surprising that under normal market conditions the spread was quite small. However, when the system came under speculative attack the capital flight could only be financed through the draft market and short of massive intervention on the part of the Central Reserve Bank, the escape valve theory would predict that the rates would diverge. This would be particularly true if one made the plausible assumption that in the short run the arbitrage mechanisms could not respond quickly enough and in sufficient volume to insure the smallness of the spread. One should also add that it was not the practice of the Central Reserve Bank to systematically intervene in the draft market.

If we turn to the two Peruvian exchange crises, 1953–54 and 1957–59, we note that

⁴ A rather naive model was tested, namely, regressing the observed spread on net long-term private foreign direct investment. Using annual data for 1950–59 the result was:

$$S = 0.92 - 0.0011F \quad R^2 = -.44 \\ (0.12) \quad (0.00007)$$

where S is the mean annual spread between the two rates, expressed as a percentage, F is the annual net long-term private foreign direct investment measured in millions of dollars and the numbers in parentheses are the respective standard errors.

⁵ Further, those exports which were subject to an export tax, the government tended to systematically overvalue (causing exporters to have to later file for tax refunds). Secondly, *ad valorem* tariffs would have cut into the profits to be had from over-invoicing imports.

TABLE 2—MEAN MONTHLY CERTIFICATE EXCHANGE RATES AND SPREADS BETWEEN THE DRAFT AND CERTIFICATE RATES, COMPUTED FROM MEAN MONTHLY RATES^a

	1953		1954	
	Certificate Rate ^b	Spread (Percent)	Certificate Rate ^b	Spread (Percent)
January	15.72	0.38	21.63	2.64
February	15.90	0.44	18.73	3.63
March	15.97	0.56	19.36	2.74
April	16.47	3.10	19.20	3.18
May	16.16	0.87	19.30	3.58
June	16.13	0.93	19.40	3.30
July	16.30	0.49	19.40	2.47
August	17.30	0.87	19.39	2.01
September	17.72	1.24	19.20	1.04
October	18.14	0.77	19.00	0.95
November	18.90	1.06	19.00	0.37
December	19.89	0.35	19.00	0.32

^a Source: *Memoria y Estadística*, 1953, 1954, Superintendent of Banks; Lima, Peru

^b soles to the dollar

in neither case was the spread very large. The 1953 crisis got underway in the latter part of the third quarter and yet during the fourth quarter the differential between the rates never rose above the overall annual mean differential for the years 1950–59. It was only following the January 1954 stabilization program that the spread became sizable. This relatively large spread persisted until the Central Reserve Bank pegged the certificate rate in October of the same year. One explanation of the observed differentials is speculation. That is, the draft market served as an effective channel for siphoning off the demand of speculators who were unconvinced of the government's ability to execute effectively its announced stabilization program. An alternative is provided by two other complementary explanations. The first is that along with the stabilization program, the Bank stepped up its systematic intervention in the certificate market, adding to its appreciating tendency, while at the same time following its normal policy of allowing the draft market to seek its own equilibrium. The equilibrating process in the draft market was seriously affected by net long-term private foreign direct investment, being a negative 12.8 million dollars.

Thus in the early stages, as the crisis was

developing, the draft market did not serve as an effective means of diverting the excess demand from the certificate market since the certificate rate depreciated 25.3 percent between its low in October 1953, and its high in January 1954. It is difficult to determine whether the draft market helped the appreciation of the sol in 1954 since there is no clear evidence on the extent of the speculation carried by the draft market during that time. However, since in the first three quarters of 1954, the draft market carried a relatively smaller portion of the total exchange transactions than it had in 1953, there is some reason to assume that the speculation was not terribly large and that the Bank's intervention in the certificate market and the negative net long-term private foreign direct investment largely accounted for the spread between the rates.

Turning to the second exchange crisis, we note that during 1957 when the Bank's peg was under attack, the spread between the rates never exceeded one percentage point. Were the draft rate serving as a channel for the excess demand for dollars then one would have expected a significantly higher premium on draft exchange, particularly if one accepts the estimates of the technical mission from the International Monetary Fund

TABLE 3—MEAN MONTHLY CERTIFICATE EXCHANGE RATES AND SPREADS BETWEEN THE DRAFT AND CERTIFICATE RATES, COMPUTED FROM MEAN MONTHLY RATES*

	1957		1958		1959	
	Certificate Rate ^b	Spread (Percent)	Certificate Rate ^b	Spread (Percent)	Certificate Rate ^b	Spread (Percent)
January	19.00	0.21	20.32	2.31	25.50	0.78
February	19.00	0.26	22.40	0.94	27.32	0.84
March	19.00	0.21	22.69	1.15	27.40	0.84
April	19.00	0.32	22.68	0.31	27.14	1.07
May	19.00	0.42	22.86	0.83	27.39	0.80
June	19.00	0.37	23.70	0.72	28.67	3.73
July	19.00	0.42	24.26	0.74	27.69	0.58
August	19.00	0.42	24.45	0.53	28.20	0.28
September	19.00	1.00	24.56	1.05	27.83	0.40
October	19.00	1.00	24.94	0.36	27.70	0.07
November	19.00	0.95	24.98	2.08	27.70	0.04
December	19.00	0.89	24.49	0.57	27.70	0.04

* Source: *Memoria y Estadística*, 1957-59, Superintendent of Banks; Lima, Peru

^b soles to the dollar

which placed the speculation on the order of ten million dollars a month during the third quarter.

Further, the role of the draft market in terms of total exchange transactions was lower in 1957 than it had been in 1953, in spite of the fact that net long-term private foreign direct investment was at its high for the period 1950-1959. During the next two years, as the economy stumbled from one abortive stabilization program to another, the relative importance of the draft market, as contrasted with 1957, fell and with the exception of four separate months the spread was always less than one percent.

The net result of these two exchange crises seems to be that the dual market arrangement added little to insure the stability of the exchange rate. This is not to say that in principle such a dual exchange system could not improve the short-run stability of the primary exchange rate, nor is it to be construed in any way as a statement that the Peruvian dual market system was a failure. What is involved is a trade-off: as the escape valve qualities of a system are enhanced by one of the markets becoming increasingly restrictive, the resource distortions will increase as will the probability of putting off the adjustment of a fundamental disequi-

librium. This brings us to our first argument against the dual markets. The fact that the Peruvian system did not work to effectively stabilize the certificate exchange rate is simply another way of saying that the authorities were unwilling to pay for this "stability" in terms of the resource distortions which a more highly restricted certificate market would have necessarily induced.⁹

Aside from the misallocations brought about by a dual exchange market, there is also the possibility that exchange controls themselves will discourage foreign investment. The Central Reserve Bank of Peru recognized this when at the suspension of the certificate system in 1960, it wrote on page 4, *Boletín Mensual*, "... the maintenance of the regime of foreign exchange certificates, in the eyes of other countries, constitutes a significant manifestation of our own lack of confidence in our currency, which was an unfavorable factor in attracting investments from abroad." The Bank later echoed the same point in a September 15, 1967 statement:

⁹ The term stability is in quotes since in fact any such stability is a false stability, with the real crisis temporarily hidden behind a set of exchange controls and restrictions.

In effect, whatever restrictions—in form of licenses, prohibitions or exchange controls or any other form—would not only work counter to their expressed end, but would slow production, would retard investment and would generate a series of distortions in economic activities, with consequent adverse effects on effective national progress. [p. 8]

In summary, we started by noting that at its inception the certificate market was a form of exchange control liberalization and that its reestablishment was a movement away from exchange control liberalization. The reestablished system is formally, at least, like the earlier one and to that extent it will not cause severe problems of resource misallocation. However, it will also be limited in its ability to stabilize the exchange rate, since such stability is bought at the price of distortions in the economy. Thus if the new system is to operate as it did in the past then it will do little to insure exchange rate stability and will cost the economy some

resource misallocations and potentially might result in a smaller volume of foreign investment. If the new system is set up to guarantee that its escape valve mechanism is effective, then this can only be done at the price of increased resource misallocations and presumably further reductions in the level of foreign investment. In short, either way the economy stands to lose, hence we conclude that the system should not have been reestablished.

REFERENCES

- L. Morse, "The Peruvian Experience with Fixed and Flexible Exchange Rates: An Empirical Examination," unpublished doctoral dissertation, Univ. Minnesota 1968.
- Central Reserve Bank of Peru, letter from Governor dated January 31, 1950, to the director of the International Monetary Fund.
- , *Boletín Mensual*, Casa Nacional de Moneda, Lima, May 1960; September, 1967. (Freely translated by the author.)

Risk and the Social Rate of Discount

By MARK V. PAULY*

A step toward clearing up controversy about the proper rate of discount to use in evaluating benefit and cost streams from public investment was provided by a recent article by William Baumol. But Baumol has left us with an uncomfortable indeterminacy. He claims that there is not and cannot be a single "optimal" social rate of discount. This is because private aversion to risk causes the rate of return r' on private investment to diverge from the rate r at which individuals are willing to loan money to the government.¹ Since public and private investments are equally riskless when the operation of the law of large numbers is considered, Baumol recommends, but does not expect to see, a subsidy to private investment to offset what investors (mistakenly) perceive as increased riskiness.

I will show in this paper that the rate r at which individuals are willing to loan money to the government necessarily understates the true social cost of the investment, and so should not be used as a measure of the social rate of discount. Hence, divergence between r and r' does not necessarily indicate the absence of an optimal rate. I will also show that, if taxation-induced distortions are ignored, the "correct" rate for evaluating government investment opportunities may be equal to, or greater than, the private rate r' . In such cases, risk does not provide grounds for a differential subsidy to private investment.

* The author is assistant professor of economics at Northwestern University. Helpful comments were received from Kenneth Arrow, William Baumol, Jack Hirshleifer, Richard Kihlstrom, and Colin Wright. Research support was provided by the Health Services Research Center.

¹ Differential taxes on income from private investment, especially the corporate income tax, cause a similar divergence, but integration of the corporate with the individual income tax, which is probably desirable on other grounds, would avoid this problem. The effect of taxation on rates of return will be ignored in this paper.

The standard argument for the lesser riskiness of public investment is based on the law of large numbers. William Vickrey and Paul Samuelson, among others, have claimed that since government is able to spread its funds over a large number of investments, it can reduce the variance in the stream of income from those investments to a greater extent than can be done with the returns from private investments. Baumol claims that this argument is not strictly correct because from the viewpoint of society, either public or private investment is equally risk-free, owing to the fact that there are large numbers of both types of investment. But since investors behave *as if* private investment is differentially risky, he argues that a divergence, a risk-premium is introduced between r and r' . Baumol therefore suggests that private investment be subsidized to bring r' down to r . If such a risk-premium does indeed exist, the suggested use of various analogous private discount rates (see Hirshleifer, DeHaven, and Milliman), or of a rate which is weighted average of public and private rates (see Arnold Harberger), would amount to treating some or all of risk as a social as well as a private cost.

In fact, however, there may be no reason why investors need to treat private investment as differentially risky. Pooling of private investment is possible, and it may be possible on at least as large a scale as is pooling of public investment. In part, pooling is provided by purchase of the securities of a large diversified firm. But, it is also provided by diversification of portfolios to include the shares of many firms. Investors with relatively large amounts of funds can engage in pooling directly through purchase of diversified portfolios, and smaller investors can purchase slices of such portfolios by such devices as mutual funds or insurance annuities. The government spreads

risk by making many wholly-owned investments, and a private investor can spread risk by purchasing small shares of many investments.³ Hence, pooling via the financial markets takes place in the private sector.

It may be that private markets for risk bearing are imperfect, and that perfect diversification is impossible, so that private pooling of individual investments is imperfect. Yet, Peter Diamond has shown that security markets may be relatively effective in pooling risks. Moreover, it is also not true that the government in fact is a perfect pooler, in the sense that it provides all possible commodity options (see J. Hirshleifer and D. L. Shapiro). In terms of its own overall accounts the government may be able to count upon the operation of the law or large numbers. But, as Diamond has suggested, this is by no means the necessary case for beneficiaries of particular risky government projects.

Suppose there is perfect pooling of individual investments in both the public and private sectors. One might conclude that the rate at which individuals would be willing to loan money for diversified private investments should, in the absence of tax distortions, approximate the rate at which individuals are willing to loan money to the government. But there is another reason why the public rate will tend to be below the private rate. Government bonds are safe or certain because the government can always tax, either explicitly or implicitly by printing money, in order to pay the interest on them. Private investment, even diversified private investments of the type just discussed, does not have such a characteristic.

³ It is true, as Baumol argues, that for society the stock of investments includes both public and private investment, and that the number of such investments is a "large number." The number of private investments in any private pooling will be less than this totality, but it can be very large, so that in such a case the divergence between marginal private risk and marginal social risk is of the smallest order of smalls. Perhaps more importantly, the existence of a large number of investments, whether made by "society" or by the private sector of society, is only a *necessary*, not a *sufficient* condition for Pareto optimality in the sense of provision of all possible commodity options and the insurance of all risks.

Its diversification guards against fluctuations in the earnings of individual firms, but the return is not protected against the variability that arises from conditions which affect all of its component firms in more or less the same way. This risk, which Hirshleifer (1966) calls "social variability", arises from occurrences such as depression, war, etc. Such a national disaster (or good fortune) affects the return on almost all investment, and would cause the mean of private rates of return to fluctuate over time. If prospective lenders are risk-aversers, the rate of return on private investment would then include a "risk premium."

The return from public investment is guaranteed not to fluctuate by the taxing power of the government. Risk-averting individuals would therefore be willing to lend government money at a lower rate, and so it appears that there is a divergence between private and public costs.

This divergence is only apparent, however, because risk is also associated with government investment. Suppose government bonds finance public investment. This investment is also subject to the social variability of depression and war. A depression, if it occurred, would make public investment in waterways, human capital, or whatever, yield a low or zero return. If something did occur which made returns on all public investments decline, the interest on the bonds sold to raise the funds would still be paid, but it would be paid by general taxpayers out of their consumption or their returns, if any, from private investment. In short, public investment implies assumption by taxpayers of the social risk associated with that investment. There is no reason to suppose that the correlation of the return from public investment with "general economic and social conditions" is any less than that of diversified private investment of equal term.⁴ In any event, public investment

⁴ Arrow (1966) seems to maintain that returns on new public investments tend to be uncorrelated or negatively correlated (as in the case of flood control) with social risk. I see no basis for such an assertion. Surely many public investments do vary directly with social risk, and some private investment are negatively correlated with it.

is certainly not completely free of variability in return. This means that the cost of public investment is not just the interest cost of funds to the government, but must also include the cost of assumption of social risk by risk-averting taxpayers.

To complete the argument, we need to consider the prospective purchaser of bonds. His willingness to make funds available to the government at r percent measures his pure time preference. But the interest rate r certainly understates the cost of a public investment. The cost of that investment is the sum of the return for sacrificing present consumption by bondholders plus the incremental risk imposed upon taxpayers in general.⁴ It is this sum, not the rate of time preference alone, which optimally ought to be equated to the expected rate of return from public investment. The rate $r' > r$ at which individuals are willing to make funds available for a diversified private investment, on the other hand, already includes a risk premium for general risk.⁵ Hence, the rate at which individuals are willing to lend money to the government necessarily understates the true cost of public investment if taxpayers are risk-averters. It is not appropriate to use the riskless rate in a world risky in the large as well as the small.

⁴ It might be suggested that, insofar as bond purchasers are also taxpayers, they will consider the risk they will suffer as taxpayers when deciding whether or not to purchase bonds, and so r will reflect some of the risk associated with government investment. This is not correct, however, because in a competitive market each bond purchaser will take the amount of bonds to be sold as given, and will not think that his purchase of a bond increases the risk he bears as a taxpayer.

⁵ Contrary to Arrow (1966), if this sort of risk, which cannot be eliminated, is the source of the divergence in rates and risk, then observation that all risk to private firms has not been eliminated by the private market is not sufficient to prove that the private market is inefficient in pooling of individual (not social) variability risk. Nor does the fact that the return on first quality corporate bonds approximates the government rate indicate that private pooling of individual investment is less perfect than public pooling. The relevant rate is the the return on private investment. The seniority of corporate debt protects the return on that security from a large part of individual and social risk, but this does not reflect any characteristic of the return on private real investment.

Baumol regards imperfect pooling of private investment to be (in a tax-free world) the source of the divergence between r and r' . I contend that even with perfect pooling of individual private investments social variability would still cause a divergence to persist. The returns from government investment are subject to much the same social variability, even though the return from money individuals lend to the government is not. If private and public individual-investment pooling are perfect, as has been assumed, then the expected divergence between observed rates does not provide grounds for a differential subsidy to private investment, since the divergence only reflects understatement of the social cost of public investment. Instead, in such a case no subsidy is warranted, and the "correct" rate of discount, for both public and private investment, is the private rate r' (or more precisely, a series of various private rates corresponding to the risk-class of the investment).⁶ It is important to note that these same conclusions also follow if private pooling is imperfect, in the sense of not providing all possible commodity options, but is no more imperfect than public pooling and provision through the political process.

Of course, the assumption of equal efficiency of public and private pooling is not the only assumption possible. If, as seems likely, not all of the risk can be eliminated from the benefit streams of government investment, then we can be sure that the "correct" rate of discount lies above the government rate r . If the government is more efficient than the private market at pooling and at providing commodity options, but is not perfect (either because of social risk or because political decision costs preclude provision of some commodity options), then the appropriate rate, *which will not be observable*, will be above the rate at which individuals are willing to loan money to the government but below the private rate. A subsidy to private investment would be warranted in such a case. If, however, the government is less efficient at

⁶ I am indebted to Hirshleifer for this point.

pooling than is the private market, then the appropriate discount rate for public investment returns would be above the private rate and a subsidy to private investment would be unwarranted. "Efficiency" in this sense is ultimately an empirical question, depending upon the number of investments made, the correlation of their returns with each other and with social risk, and the availability of possible commodity options. Empirical research on public market and private market imperfections would be useful to help provide the answer.

REFERENCES

- K. J. Arrow, "The Role of Securities in the Optimal Allocation of Risk-Bearing," *Rev. Econ. Stud.*, Apr. 1964, 31, 91-96.
- , "Discounting and Public Investment Criteria," in A. V. Kneese and S. C. Smith, eds., *Water Research*, Baltimore 1966.
- W. J. Baumol, "On the Social Rate of Discount," *Amer. Econ. Rev.*, Sept. 1958, 58, 788-802.
- P. A. Diamond, "The Role of a Stock Market in a General Equilibrium Model with Technological Uncertainty," *Amer. Econ. Rev.*, Sept. 1967, 57, 759-76.
- A. Harberger, "The Social Opportunity Cost of Capital: A New Approach." Paper presented at the Annual Meeting of the Water Resources Research Committee, Western Agricultural Research Council, Denver, Dec. 11, 1968.
- J. Hirshleifer, "Investment Decision under Uncertainty," *Quart. J. Econ.*, May 1966, 80, 252-76.
- , J. C. DeHaven, and J. W. Milliman, *Water Supply: Economics, Technology, and Policy*. Chicago 1960.
- and D. L. Shapiro, "The Treatment of Risk and Uncertainty," paper prepared for Joint Economic Committee, *Economic Analysis of Public Expenditure Decision*, 90th Cong., 2nd Sess., Washington 1969.
- P. A. Samuelson, "Principles of Efficiency—Discussion," *Amer. Econ. Rev. Proc.*, May 1964, 54, 93-96.
- W. Vickrey, "Principles of Efficiency—Discussion," *Amer. Econ. Rev. Proc.*, May 1964, 54, 88-92.

A Nonlinear Test of the Random-Walk Hypothesis

By R. L. CROUCH*

Academic economists have accumulated a mountain of evidence in the last fifteen years to the effect that successive changes in stock prices follow a random-walk process.¹ In fact, it is probably no exaggeration to say that the random-walk hypothesis is one of the best, if not the best corroborated economic hypothesis that we have. However, as Eugene Fama has pointed out, the evidence which has been used to corroborate the random-walk hypothesis, although strong, does suffer certain drawbacks. In particular, he states that

... there are also possible shortcomings to the serial correlation and runs tests from a statistical point of view. For example, both of these models only test for dependence which is present all through the data. It is possible, however, that price changes are dependent only in special conditions. For example, although small changes may be independent, large changes may tend to be followed consistently by large changes of the same sign, or perhaps by large changes of the opposite sign. [p. 80]

By comparison of distributions of all price changes with distributions of successive price changes ("successive" to large price changes), Fama concludes that "... there is some evidence that large changes tend to be followed by large changes of either sign, ..." (p. 87). Fama arbitrarily defined

"a large daily price change... as a change in log price greater than 0.03 in absolute value" (p. 86).

I think that Fama was on the right track here. My own test is an attempt to capitalize on his initiative. Section I rationalizes a technique for isolating "special conditions" in the securities market which is somewhat less arbitrary than Fama's method, and Section II tests for intertemporal dependence in successive price changes in such circumstances.

I. Identifying Special Conditions

Visualize the market for a security. In the short run, the existing supply is fixed; the supply curve is perfectly inelastic. The total demand curve intersects the supply curve and the two jointly determine the equilibrium price. The demand and supply curves are of the stock (cf. flow) variety; they represent the demand to hold and the supply that must be held. As such, they have a time *reference* but no time *dimension*.

When total demand is constant, so, also, will price be constant. However, the individual demands which comprise the total demand may be undergoing offsetting changes. This immediately implies that transactions may certainly occur without there being any change in price of the security. Define the volume of transactions which occur as a result of the mutually offsetting switches in individual demands as *normal volume*. Assume, now, that there is an increase in total demand (due, say, to some piece of new fundamental information concerning the security). The demand curve shifts to the right. At the prevailing price, excess demand exists. The increased demand manifests itself in increased transactions. What may be described as *abnormal volume* appears in the market. The excess demand is eliminated by a rise in price. At the new, higher equilibrium price, *normal volume* will

* The author is associate professor of economics at the University of California, Santa Barbara. This paper comprises a section of a larger paper presented at the Third Far Eastern Meeting of the Econometric Society, Tokyo 1968. I wish to thank my discussant, T. Hatanaka, for his penetrating criticism in general and, especially, for guiding me out of one blind alley. Robert Clower also provided helpful comments.

¹ Although the random-walk hypothesis is relatively ancient, the corroboratory evidence is almost exclusively a product of the last fifteen years, beginning with Maurice Kendall's work. A good compendium on the random-walk hypothesis theory is contained in the book edited by Paul Cootner.

TABLE 1—DAILY VOLUME AND THE ABSOLUTE VALUE OF PRICE CHANGES

SECURITY	Result	S.E.E.*	R ²	Split, Date	Time Interval, N
Anaconda	$V = 83.0 + 66.3 \dot{p} $ (19.4) (17.1)	104.0	.208	—	1/2/63–6/9/67 1117
United Airlines	$V = 104.0 + 12.5 \dot{p} + 161.9d$ (23.8) (6.9) (19.6)	120.0	.299	2–1 5/9/66	1/2/63–6/30/67 1133
American Airlines	$V = 146.0 + 18.5 \dot{p} + 226.1d$ (37.6) (7.6) (7.2)	114.0	.112	2–1 6/13/67	1/2/63–6/30/67 1133
United Aircraft	$V = 50.0 + 19.8 \dot{p} + 108.1d$ (13.3) (8.5) (17.7)	95.0	.317	3–2 11/2/65	1/2/63–6/30/67 1133
U.S. Steel	$V = 216.0 + 215.2 \dot{p} $ (25.1) (17.7)	206.0	.216	—	1/2/63–6/30/67 1133

* Standard error of estimate.

Source: *The Wall Street Journal*.

once more prevail. If, on the other hand, there had been a decrease in total demand, then, when the demand curve shifted left, an excess supply would exist at the original price. Increased transactions, *abnormal volume*, would again ensue as unwilling holders attempt to lighten their portfolio. The excess supply is eliminated by a decline in price. At the new, lower price, *normal volume* will emerge again. It will be apparent that an implication of this model is that the volume of transactions V and the *absolute* value of the change in the price of a security $|\dot{p}|$ are related. Stochastically, and linearly, we postulate:

$$(1) \quad V_t = a_0 + a_1|\dot{p}|_t + a_2d_t + u_t$$

Where the variables, with the exception of d , are as defined above. The variable d is a dummy designed to capture the effect of stock-splits. A stock-split will certainly affect volume systematically; consequently, such an event cannot be relegated to the error term u . In practice, d is assigned the value zero prior to a stock-split and unity after a stock-split (if any).

Equation (1) was estimated for several individual securities.² The results are re-

ported in Table 1. All the coefficients are significant as revealed by the t -ratios reported in parentheses in the second row of each result. The smallest of these ratios is 7.2. Normal volume can be calculated from these results by setting $|\dot{p}|$ equal to zero. The pre- and post-split normal volume for each security is reported in Table 2 along with the pre- and post-split abnormal volume. The latter is calculated by adding two standard errors of estimate to normal volume. These estimates of abnormal volume are those used to identify special conditions in the market for a particular security.

II. Serial Correlation In Special Conditions

The next step is to determine whether or not stock price changes under such special conditions are nonrandom. The specific technique employed was to comb the complete sample for each security for observations involving abnormal volume. The daily price change associated with abnormal volume was then serially correlated with successive daily price changes for lags of from one to ten days. The serial correlation co-

is essential to work with such a large number of original observations in order to be left with a viable number of special condition observations. For example, the eleven hundred-plus original observations typically yield a mere sixty or so special condition situations.

² The daily data on volume and price, in excess of eleven hundred observations for each stock, were kindly made available to me by Robert R. Rooney. It

TABLE 2—NORMAL AND ABNORMAL VOLUME
(hundreds of shares)

SECURITY	Normal Volume		Abnormal Volume	
	Pre-Split	Post-Split	Pre-Split	Post-Split
Anaconda	83	—	291	—
United Airlines	104	266	344	506
American Airlines	146	372	374	600
United Aircraft	50	158	240	348
U.S. Steel	216	—	629	—

efficients obtained for each security when this procedure was followed are reported in Table 3. In the present context, serial correlation amounts to a nonlinear test. Those serial correlation coefficients that are significant at the 5 percent level are marked with an asterisk. There does not appear to be any intersecurity uniformity as to which serial correlation coefficients are, and which are not, significant; they occur at different lags and have different signs. On this evidence, there is little or no predictable and persistent pattern of dependence in the successive changes of stock prices even under special conditions. Moreover, if one is persuaded, as I am, by the evidence that Benoit Mandelbrot and Fama have marshalled to the effect that stock price changes possess stable Paretian distributions, then

the normal procedure for estimating the standard errors of the serial correlation coefficients *underestimates* the true variability of the coefficient; (see Fama pp. 69–70). In this case, the significance of the serial correlation coefficients marked with an asterisk is still subject to doubt. Side stepping this qualification, however, what is the prospect open to someone trying to capitalize on the best evidence of dependence in successive stock price changes uncovered by the above results? The one-day lag in United Aircraft has the largest serial correlation coefficient. A person planning to exploit this situation would wait for special conditions to arise in the market for this security (in the form of abnormally high volume), and then follow the market for one day. That is to say, if the price went up when special conditions emerge, he would buy at the next day's opening price and sell at the following day's opening price: if price went down, he would go short at the next day's opening price and cover at the following day's opening price. Such a strategy would pauperize the speculator and enrich his broker. It is perfectly designed to transfer the speculator's wealth to his broker in the (not so) long run. Consider the following regression result:

$$\hat{p}_{t+1} = 0.102 + 0.22\hat{p}_t; \\ (.455) \quad (2.56)$$

$$R = .360, \quad R^2 = .130.$$

This indicates that, under special conditions, a \$1.00 change in price today is followed by a \$.022 change in price tomorrow in the same

TABLE 3—SERIAL CORRELATION COEFFICIENTS UNDER SPECIAL CONDITIONS FOR
Lag $\theta = 1, 2, \dots, 10$ (DAYS)

SECURITY	LAG										
	1	2	3	4	5	6	7	8	9	10	N
Anaconda	-.012	-.173	-.117	+.051	+.155	+.147	-.165	+.010	-.121	-.297*	82
United Airlines	+.131	-.140	-.028	+.181	+.202	-.057	-.084	+.291*	-.049	-.021	50
American Airlines	+.053	+.263*	+.126	-.070	+.269*	+.027	-.115	-.019	-.310*	+.092	65
United Aircraft	+.360*	-.199	+.010	+.121	+.097	-.125	-.021	-.177	-.215	-.161	46
U.S. Steel	-.312*	-.039	-.151	-.239*	+.164	-.051	-.123	-.188	+.072	+.082	76

direction. This is not enough to cover transactions costs by a long way. (During this period United Aircraft had an average price around \$75.00.) If an investor bought one round lot, his expected gross return would be \$22.00; his transactions costs for the purchase and sale of one round lot would be \$93.00.³ Thus, on each round trip, he would make a net loss of \$71.00. The conclusion is inevitable. Whatever dependence does exist, even under special conditions, when the market is adjusting to major

³ The one-way transactions cost of one round lot of stock selling in the \$75.00 range is equal to "\$39.00 plus one-tenth the price of one share," i.e., \$39.00 plus \$7.50 equals \$46.50. Thus, the round trip cost is \$93.00 (excluding a small transfer tax on the purchase).

changes, it is too small to be exploited profitably.

REFERENCES

- P. H. Cootner, ed., *The Random Character of Stock Market Prices*, Cambridge, Mass. 1964.
- E. F. Fama, "The Behavior of Stock Market Prices," *J. Bus., Univ. Chicago*, Jan. 1965, 38, 34-105.
- M. G. Kendall, "The Analysis of Economic Time-Series," *J. Roy. Statist. Soc., Part I* 1953, 116, 11-25.
- B. Mandelbrot, "The Variation of Certain Speculative Prices," *J. Bus., Univ. Chicago*, Oct. 1963, 36, 394-419.

IS, LM, and External Equilibrium: A Graphical Analysis

By DWAYNE WRIGHTSMAN*

The *IS-LM* exposition of general equilibrium in the domestic money and goods markets excludes the problem of balance of payments equilibrium when the economy engages in foreign trade under conditions of fixed foreign exchange rates.¹ This expositional shortcoming is easily rectified by superimposing an external equilibrium condition onto the *IS-LM* framework. A new external equilibrium curve, called the *EE* curve, is derived geometrically as shown in the four-section diagram of Figure 1.

In the lower right-hand section of Figure 1, net exports of goods and services is shown to vary inversely with the level of national income. A change in income is expected to change demand for imports in the same direction without having any appreciable effect on demand for exports. At low levels of income, the level of net exports is positive, meaning that exports exceed imports. But at high income levels, the level of net exports is negative indicating imports in excess of exports.

In the upper left-hand section, net foreign investment is shown to vary inversely with the domestic interest rate. A change in the interest rate is expected to change the domestic economy's investment of funds in foreign economies in the opposite direction, and foreign economies' investment of funds in the domestic economy in the same direction. The net foreign investment section of

the diagram allows for both positive and negative levels of net foreign investment. At low interest rates, domestic investors will be net buyers of foreign securities. But if the interest rate is high enough, foreigners will be net buyers of domestic securities, so that net foreign investment will be negative.

In the lower left-hand section, the 45-degree line represents equality between net exports of goods and services and net foreign investment, which, abstracting from problem of unilateral transfers, and net factor income, is the condition for external equilibrium (meaning no surplus or deficit in the balance of payments). The *EE* curve in the upper right-hand section of Figure 1 is subsequently derived by geometry from the two behavioral functions and the equilibrium condition graphed in the other three sections. It contains combinations of the interest rate level and the level of income which provide equilibrium in the balance of payments.

The *EE* curve has a positive slope as does the *LM* curve. A comparison of the steepness of the slopes is thus in order. Is *EE* steeper or less steep than *LM*? And do the two curves intersect?

The absolute slope of the *EE* curve is the ratio between the responsiveness of net exports of goods and services to the level of income and the responsiveness of net foreign investment to the rate of interest. The absolute slope of the *LM* curve, on the other hand, is the ratio between the responsiveness of the transactions demand for money to the level of income and the responsiveness of the asset demand for money to the rate of interest. There is no a priori basis for generalizing that one of these ratios exceeds the other for all economies at all times. However, a case can be made for the proposition that the slope of the *EE* curve is normally steeper

*The author is associate professor of finance at the University of New Hampshire. He is indebted to William Hosek, James Schulz, Richard Whyte, and the managing editor for their helpful comments and suggestions.

¹ An *IS* curve is a curve showing combinations of the level of national income and the level of interest rates which generate equilibrium in the market for goods and services. An *LM* curve does the same for the money market. For background on the nature of these curves, see John R. Hicks.

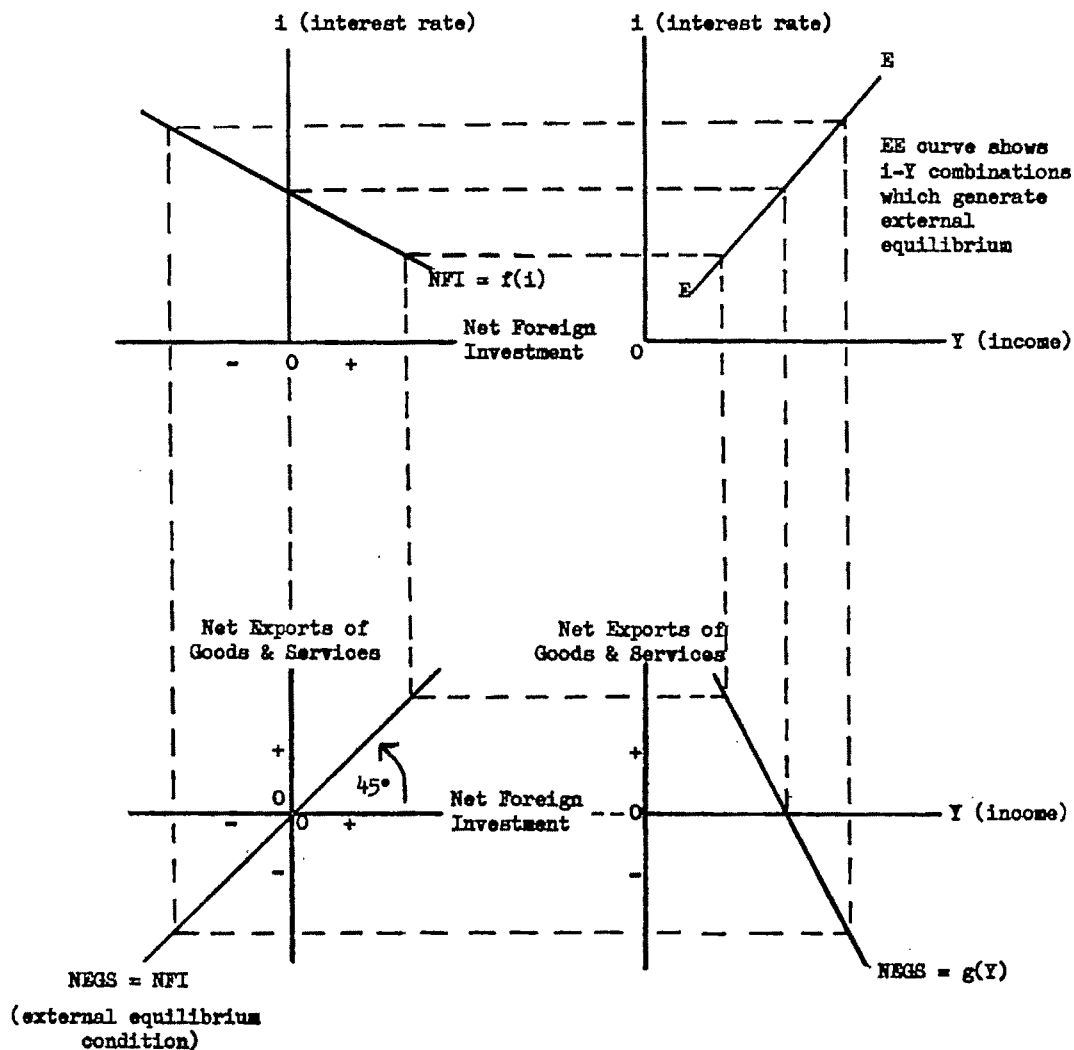


FIGURE 1

than the slope of the LM curve and that it (EE) normally intersects LM from below (moving to the right).

First, consider an EE - LM intersection point at some positive i and Y . This merely means that there is some interest rate above zero and some level of income above zero capable of ensuring simultaneous equilibrium in the domestic money market and the balance of payments. Surely this is the normal circumstance. Countries occasionally do approach such an equilibrium, and often those that do not, could, were it not for pressing internal considerations, for

example, maintenance of high-level employment.

Now consider what would happen to an economy initially in equilibrium if a general recession were to strike (without its being anticipated by fiscal and monetary authorities). Income and the interest rate would fall. In a financially mobilized country such as the United States, money demanders would be expected to adjust their cash balances rapidly so that the lower i - Y combination would still be on the LM curve, only lower down. Thus the money market would soon be back in equilibrium. But

what would happen to the external equilibrium in the balance of payments? One would expect, normally, that a general recession would disrupt external equilibrium and create a payments surplus; that is, the recession's income effect of increasing net exports of goods and services would outweigh its interest rate effect of increasing net foreign investment. If this is the case, it means that the new interest rate is too high and the new income too low to prevent a payments surplus. The *EE* curve, consequently, lies below the *LM* curve for income values to the left of the *EE-LM* intersection. (In the case of an unanticipated boom accompanied by rapid adjustment in money demand, one would expect a payments deficit to develop, meaning that *EE* lies above *LM* for *Y* values to the right of *EE-LM* equality.) Thus, under these circumstances, which are assumed to be normal, the relation between the *EE* and *LM* curves is that they intersect with the former being steeper than the latter.

Unlike the *LM* and *EE* curves, the *IS* curve has a negative slope. However, the *IS* curve has a steeper negative slope, *ceteris paribus*, in an open economy than in a closed economy. A definition of the *IS* curve will help to explain why this is so. The *IS* curve depicts all *i-Y* combinations which are consistent with a zero excess demand for goods and services. For the closed economy, this implies that there is *ex ante* equality between saving and domestic investment and that the absolute slope of the *IS* curve is the ratio between the responsiveness of saving to the level of income and the responsiveness of domestic investment to the rate of interest. For the open economy, it implies that saving minus net exports of goods and services equals domestic investment and that the slope of *IS* is the ratio between the responsiveness of saving minus net exports to the level of income and the responsiveness of domestic investment to the rate of interest. Since the response of net exports of goods and services to the level of income is negative, it follows, *ceteris paribus*, that the absolute ratio between the responses in the open economy is greater than that for the closed economy. However, in both cases the responses have opposite signs so that *IS* is

negatively sloped regardless of whether or not the economy is open or closed.

With this as background, it is now appropriate to consider the interrelations of the *IS*, *LM*, and *EE* curves. Full equilibrium in the money market, the goods market, and the balance of payments is satisfied only (but not necessarily) when the three curves pass through a common point. If they do not, the resulting disequilibrium is likely to be concentrated most in that market which bears the least influence on the rate of interest and the level of income. In the United States the least influential of these three forces is the balance of payments. Thus, *EE* is not apt to be as much income and interest rate determining as are *IS* and *LM*. The point toward which *i* and *Y* are likely to be most strongly attracted is the point of intersection between the *IS* and *LM* curves.

Passage through a common point is a necessary but not a sufficient condition for full equilibrium. If the common point lies beyond the full-employment income limit on the income axis, there will be disequilibrium in the form of inflation because the equilibrium income will be unobtainable; that is, aggregate demand for goods and services will exceed capacity supply. On the other hand, if the common point lies inside the full-employment income limit, the income disequilibrium problem vanishes but an equally distasteful problem, namely unemployment, emerges.

Internal and external problems can be analyzed separately in the *IS*, *LM*, and *EE* framework, but it is more interesting and instructive to consider them together. Consider, for example, the dual-problem case of unemployment coupled with a payments deficit. This is the case that gave the United States much trouble in the late 1950's and early 1960's and could conceivably re-emerge if and when the Viet Nam conflict ends. In Figure 2, *IS* intersects *LM* below and to the right of the *EE* curve with the result that joint monetary and product equilibrium produces an income that is too high and an interest rate that is too low to prevent a balance of payments deficit. At the same time, *IS* intersects *LM* to the left of the *F*-line (the full-employment income

limit), producing an income which is too low for full employment.

In this case, if the monetary and fiscal authorities do nothing, the problem of the payments deficit will solve itself because gold outflows (in the absence of defensive open market operations) will decrease the money supply, shifting the LM curve up and to the left until it intersects the EE and IS curves. Realistically, the monetary authority is not likely to let this happen (at least not in the United States) because this would make the unemployment problem worse. Nevertheless, monetary contraction is superior to contractionary fiscal policy for the purpose of restoring external equilibrium. Compared to contractionary fiscal policy (IS shifted down and to the left), contractionary monetary policy (LM shifted up and to the left) can restore external equilibrium at a smaller cost in terms of lost income. In Figure 3, BC is the income lost when contractionary monetary policy is used to restore external equilibrium. When contractionary fiscal policy is used for this purpose, the income lost is AC , or AB plus BC . The net amount of income saved by using contractionary monetary policy over contractionary fiscal policy is, then, the segment AB .

Suppose, however, that full employment, not external equilibrium, is the primary

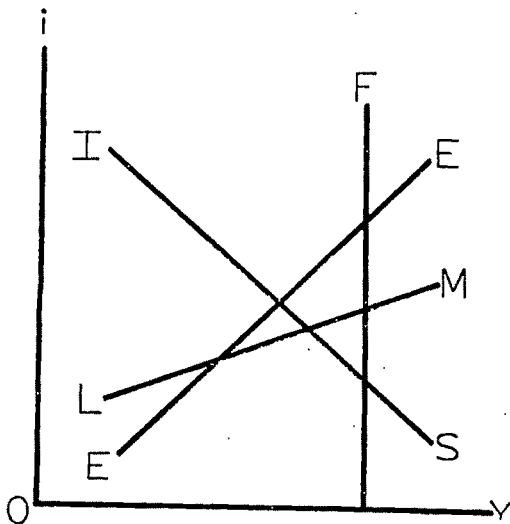


FIGURE 2 UNEMPLOYMENT-DEFICIT CASE

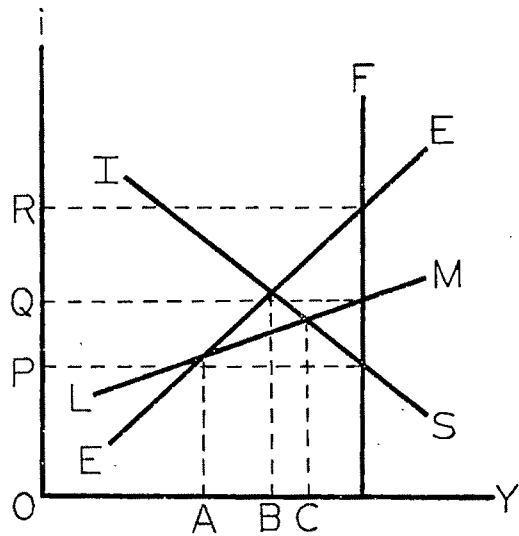


FIGURE 3

goal. For this purpose, it turns out that monetary policy is not as good a tool as fiscal policy. Expansionary monetary policy (LM shifted down and to the right) restores full employment at a greater cost (in terms of a greater payments deficit) than expansionary fiscal policy (IS shifted up and to the right). In Figure 3, QR is the deficiency in the interest rate level required for external equilibrium when expansionary fiscal policy is used to restore full employment. PR , or PQ plus QR , is the interest rate deficiency when expansionary monetary policy is used for this purpose. The net deficiency in the interest rate required to provide external equilibrium at full employment avoided by using expansionary fiscal policy over expansionary monetary policy is, consequently, the segment PQ .

The analysis thus shows that monetary policy has a comparative advantage over fiscal policy for the purpose of meeting the external equilibrium objective and that fiscal policy has a comparative advantage over monetary policy for meeting the full employment goal. The significance of the comparative advantage relations between monetary and fiscal policy is that the two problems of unemployment and a payments deficit can be simultaneously resolved by applying a mix of contractionary monetary

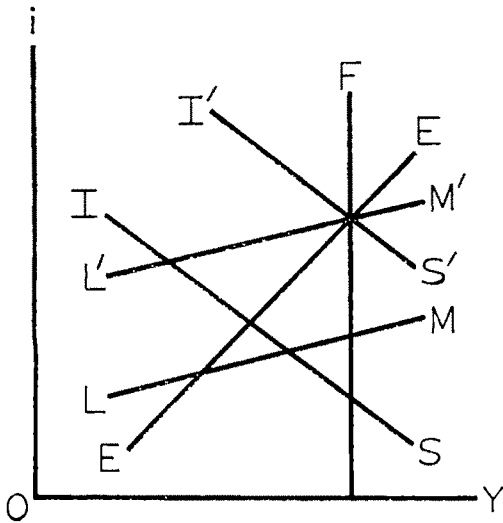


FIGURE 4

policy and expansionary fiscal policy.² In Figure 4, contractionary monetary policy is applied shifting the LM curve up and to the left to become $L'M'$. Expansionary fiscal policy is applied at the same time shifting the IS curve up and to the right to become $I'S'$. Where EE , $L'M'$, and $I'S'$ converge at the F -line the economy is at full employment with external equilibrium. Full employment income is obtained because a sufficient amount of expansionary fiscal policy is applied. External equilibrium is achieved because the expansionary fiscal policy and the contractionary monetary policy raise the interest rate to the necessary level. Of course, if an economy is not willing for various reasons to subject itself to higher interest rates, the solution just described is merely academic.³

A second interesting, dual-problem case to analyze is that of inflation coupled with a payments surplus. This case is the converse of the unemployment-deficit case above and suggests the converse solution of applying

² Such a mix is in accordance with Robert Mundell's principle of matching policy instruments with policy objectives on the basis of comparative advantage.

³ For a discussion of this and related problems associated with the monetary and fiscal policy mix, see Richard Ablin. Another treatment of a number of misgivings about the solution is found in Anne O. Krueger.

a mix of easy monetary policy and restrictionist fiscal policy. However, a country facing inflation and a surplus may not need to resort to policy measures as it does in the unemployment-deficit case because the two cases are asymmetrical. In the unemployment-deficit case, income is not out of equilibrium despite the existence of unemployment if there is wage inflexibility in the labor market. In this situation, government intervention is used as a substitute for market equilibrating forces to bring the unemployment to an end. In the inflation-surplus case, on the other hand, income is necessarily out of equilibrium and continues to be as long as the inflation persists. In this case, the rising price level generates equilibrating forces which work to end the payments surplus and the inflation without recourse to monetary and fiscal policy.

What is the nature of these equilibrating forces? When a country is confronted with inflation and a balance of payments surplus, its IS , LM , and EE curves are positioned in a way that makes all three curves shift. In Figure 5, IS , LM , and EE are positioned to indicate inflation and a payments surplus. The intersection of IS and LM lies to the left of the EE curve but to the right of the F -line (the full-employment income limit). The equilibrium level of income correspond-

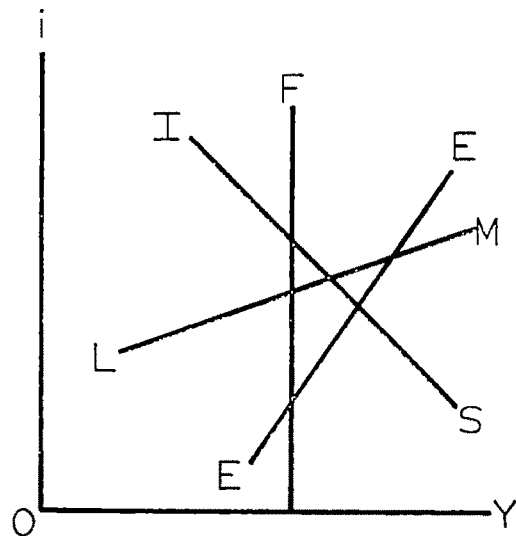


FIGURE 5 INFLATION-SURPLUS CASE

ing to the $IS-LM$ intersection is, consequently, too high to prevent inflation but too low to prevent a payments surplus. However, the positioning of the three curves is not lasting. The LM curve is potentially affected in three ways. First, the payments surplus generates an inflow of gold which, in the absence of defensive open market operations, increases the money supply, shifting LM to the right. Second, the rising absolute price level causes the transactions demand for nominal cash balances to increase, which shifts the LM curve to the left. And third, according to Friedman, the rate of inflation, if increasing, reduces the demand for money, shifting LM to the right, and, if decreasing, it increases money demand, shifting LM to the left. On balance, LM will shift to the right or to the left, depending on the relative strength of these three effects.

Inflation can also cause the IS and EE curves to shift. First, IS is shifted to the left if the inflation, working through the real balance effect, decreases consumption. The real balance effect thus serves, in the absence of elastic price expectations, to make inflation self-defeating. Second, EE should shift to the left because inflation decreases net exports of goods and services for given values of real income. A decrease in net exports, in turn, shrinks the payments surplus.

The IS curve is further shifted because of the now shrinking payments surplus. With inflation causing net exports of goods and services to decrease for given values of real income, IS has to shift to the left to remain consistent with the zero excess demand condition that saving minus net exports equals domestic investment. However, if net exports are small relative to total national product, as they are in the United States, this leftward shift in IS will not be great.

The shifts in IS , LM , and EE are generally in the direction of achieving equilibrium. There is, however, a timing consideration in reaching equilibrium. The inflation cannot be brought to an end before the payments surplus is reduced to zero because as long as there is a surplus there will be an inflow of gold and an increase in the money supply

(again, assuming no defensive open market operations) which will keep the inflation alive. If the inflation is still healthy when the payments surplus is approaching zero, the surplus is apt to turn into a deficit. The combination of inflation and a deficit will shift the LM curve to the left due to the price level effect on the demand for money and the gold inflow effect on the supply of money. This shift will serve to reduce both the inflation and the deficit, but it will not reduce the deficit to zero before it eliminates the inflation because as long as there is any inflation, there will be a decrease in net exports of goods and services which will keep the deficit alive. If the deficit is of any size when the inflation is ended, the money supply will continue to contract (the LM curve will continue to shift to the left) until external equilibrium is restored. In the meantime, however, equilibrium income will fall below the full-employment income limit producing unemployment. Thus, there is the possibility that the inflation-surplus problem can turn, by itself, into the unemployment-deficit problem discussed earlier. Should this happen, government can intervene to apply the appropriate mix of monetary and fiscal policy.

REFERENCES

- R. Ablin, "Fiscal-Monetary Mix: A Haven for the Fixed Exchange Rate?" *Nat. Banking Rev.*, Dec. 1966, 4, 199-204.
- M. Friedman, "The Supply of Money and Changes in Prices and Output," *The Relationship of Prices to Economic Stability and Growth: Compendium of Papers Submitted by Panelists*, Joint Economic Committee, 85th Cong., 2nd Sess., Washington 1958, 241-256.
- J. R. Hicks, "Mr. Keynes and the 'Classics': A Suggested Interpretation," *Econometrica*, Apr. 1937, 5, 147-159.
- A. O. Krueger, "Balance of Payments Theory," *J. Econ. Lit.*, Mar. 1969, 7, 1-26.
- R. A. Mundell, "The Appropriate Use of Monetary and Fiscal Policy for Internal and External Stability," *International Monet. Fund Staff Pap.*, Mar. 1962, 9, 70-79.

Reserve Policies of Central Banks: Comment

By JOHN H. MAKIN*

Helmut A. Hagemann has provided a useful analysis of reserve policies of central banks in his recent article in this *Review*. However, I am afraid that the unavailability of data has led him into wishful thinking about the value of the dependent variables in his exercise.

I refer specifically to the use of $(F/R)_i$, foreign exchange of country i , relative to total reserves of country i as a proxy for $(D/R)_i$;¹ dollar-denominated foreign exchange of country i , relative to total reserves of country i .

Let us consider what distortions may arise from using $(F/R)_i$ where $(D/R)_i$ ought to be.

Let:

$$(1) \quad F_{it} = D_{it} + O_{it}$$

$$(2) \quad R_{it} = D_{it} + O_{it} + G_{it}$$

Where:

F_{it} = foreign exchange holdings in country i at time t .

D_{it} = dollar² holdings in country i at time t .

O_{it} = nondollar foreign exchange holdings in country i at time t .

G_{it} = gold holdings in country i at time t .

R_{it} = total reserves in country i at time t .

Define the ratio Q .

$$(3) \quad Q_{it} = F_{it}/R_{it}$$

$$(4) \quad \log Q_{it} = \log F_{it} - \log R_{it}$$

* Department of economics, University of Wisconsin, Milwaukee.

¹ In his article, Hagemann in fact uses D where F ought to be. I intend to show that the good results hoped for by wishing $D=F$ may be quite misleading.

² Since the IMF's International Financial Statistics define foreign exchange as holdings by monetary authorities of bank deposits, Treasury bills, and short- and long-term government securities denominated in convertible currencies, D_{it} is taken to include all dollar-denominated items of this sort in country i at time t , except long-term dollar assets which were nearly zero for the countries and for the time period being examined.

Substituting into (4) from (1), (2), and (3) and differentiating with respect to time gives:

$$\begin{aligned} \frac{d \log Q_i}{dt} &= (\alpha_i - \phi_i) \frac{d \log D_i}{dt} \\ &+ (\beta_i - \lambda_i) \frac{d \log O_i}{dt} \\ &- \mu_i \frac{d \log G_i}{dt} \end{aligned} \quad (5)$$

Where for country i :

$$\alpha_i = D_i/F_i \quad \phi_i = D_i/R_i$$

$$\beta_i = O_i/F_i \quad \lambda_i = O_i/R_i$$

$$\mu_i = G_i/R_i$$

$$\alpha_i + \beta_i = 1 \quad \phi_i + \lambda_i + \mu_i = 1$$

Equation (5) tells us that the percentage rate of change of $(F/R)_{it} = Q_{it}$ depends on the percentage rate of change of dollars, other foreign exchange and gold held by country i as well as upon the relative shares of dollars and other currencies in foreign exchange and the relative share of dollars, other currencies, and gold in total reserves.

Note that $D_i = F_i$, as Hagemann effectively assumes implies $\beta_i = \lambda_i = 0$ and changes in D_i have an effect of the same sign on Q_i , unless outweighed by changes in gold holdings.³

³ Bear in mind that changes in dollar holdings and in gold holdings may be related. Suppose gold is added to reserves of country i for some of the dollars given up. Then:

$$dG_i/dt = -a_i dD_i/dt \quad (0 < a_i < 1)$$

If $R_i = D_i + G_i$ with $O_i = \text{zero}$, then $Q_i = D_i/R_i$

And:

$$\frac{d \log Q_i}{dt} = (1 - \phi_i) \frac{d \log D_i}{dt} + a_i/R_i \frac{d D_i}{dt}$$

The rate of change of Q_i is sensitive to ϕ_i , the share of

I have computed the values of α_i , ϕ_i , β_i , λ_i , and μ_i for the countries⁴ which Hagemann has examined. Since these values are likely to change over time, (equation (5) is valid only at an instant in time with the ratios taken as constant), I have calculated them at two points in time: 1957, fourth quarter; 1962, first quarter. I believe these are periods within the time span of Hagemann's investigation.⁵

TABLE 1

Country ^a	1957-IV		1962-I	
	($\alpha-\phi$)	($\beta-\lambda$)	($\alpha-\phi$)	($\beta-\lambda$)
1	.26	.48	.59	.32
2	.66	.18	.63	.08
3	.19	.03	.21	.22
4	.08	.72	.36	.44
5	.32	.30	.31	.30
6	.28	.23	.57	.06
7	.43	.52	.12	.25
8	.39	.10	.16	.16
9	.33	.0	.49	.17
10	.01	.02	.27	.0

^a Numbers are used to represent the countries of Hagemann's study in random order.

Now we return to equation (5) and ask the question: In how many cases, might the behavior⁶ of $(D/R)_i$ over time be erroneously inferred from the behavior of $(F/R)_i$ over time?

Ignoring, for the time being, the behavior of gold and given 20 cases, (10 countries in two time periods), there are 11 cases where the change in O_i must be assumed to have the same sign as the change in D_i in order that the behavior of $(D/R)_i$ be inferred with certainty from the behavior of $(F/R)_i$.⁷

Consider the most extreme case: Country 4 in 1957. Rewriting (5) and substituting in

dollars in reserves of country i as well as to a_i , the degree to which dollars are exchanged for gold.

⁴ Save for Switzerland since the data are distorted by BIS holdings.

⁵ Hagemann does not make explicit the exact time periods which the data for each country cover. He does specify excluding, "... the initial observations in the early 1950's ..."

⁶ This only concerns behavior regarding the correct sign of $d \log (D/R)_i / dt$ inferred from $d \log Q_i / dt$.

⁷ In 1957: Countries 1, 4, 5, 6, 7, and 10. In 1962: Countries 3, 4, 5, 7, and 8.

values for the share parameters.

$$(5') \quad \frac{d \log Q_4}{dt} = .08 \frac{d \log D_4}{dt} + .72 \frac{d \log O_4}{dt} - .80 \frac{d \log G_4}{dt}$$

If $d \log D_4 / dt < 0$ and $d \log O_4 / dt > 0$, unless the percentage fall in dollars is more than nine times the rise in other foreign exchange (ignoring gold)⁸, the wrong inference about the sign of the change of $(D/R)_4$ is made from the sign of the change of $(F/R)_4$.

It should be noted that the behavior of dollar and gold holdings may be related when $d \log G_i / dt$ is not equal to zero.⁹

In sum, there is a strong case to be made for doubting Hagemann's results due to an identifiable error in the specification of the dependent variables in his study. The misspecification implies that inferences about the behavior of $(D/R)_i$ over time,¹⁰ from the behavior of $(F/R)_i$ can be (and are), misleading.

REFERENCE

H. A. Hagemann, "Reserve Policies of Central Banks and Their Implications for U.S. Balance of Payments Policy," *Amer. Econ. Rev.*, Mar. 1969, 59, 62-77.

⁸ In nine of the 20 cases under consideration, $d \log G_i / dt$ was zero or less than 1 percent. The average for all of the cases was 3.9 percent. (Averages are of absolute values.)

⁹ Recalling fn. 3 where $d G_i / dt = -a_i d D_i / dt$, equation (5) becomes

$$\frac{d \log Q_i}{dt} = (\alpha_i - \phi_i) \frac{d \log D_i}{dt} + (\beta_i - \lambda_i) \frac{d \log O_i}{dt} + a_i / R_i \frac{d D_i}{dt}$$

If D_i is falling, (G_i rising), then the last term aids the net effect and the term $(\beta_i - \lambda_i) d \log O_i / dt$ would have to outweigh it for the inferred sign of $d \log (D_i/R_i) / dt$ to be incorrect.

In any case, in all of the examples considered, the degree of $d \log D_i / R_i / dt$ inferred from $d \log Q_i / dt$ is distorted.

¹⁰ The question arises as to distortions in the cross-sectional analysis. The move through space is analytically equivalent to a move through time. The likelihood that shares can be held constant as at an instant in time falls, however, as the continuum over country space is not as smooth as the continuum through time.

Reserve Policies of Central Banks: Reply

By HELMUT A. HAGEMANN

The comment of John Makin gives me an opportunity to elaborate on some of the implications the existence of a second reserve currency might have in the model and the empirical tests.

In the model in Part I, we have assumed that there are only two reserve assets, gold and one reserve currency (dollars), and investigated what economic factors might determine the reserve policies of central banks. Of these factors, some characterize the country in question, others characterize the reserve-currency country (United States), and again, others the relation between them.

If a second reserve currency (sterling) is introduced into the model, additional variables come into play which reflect the economic position of the second reserve-currency country (United Kingdom), its relation to the specific country as well as its relative position to the first reserve-currency country. Thus, if a country can hold pounds in addition to gold and dollars, it might rearrange its portfolio and substitute pounds for gold and/or dollars.

Under the present international monetary system, pounds cannot be converted directly into gold, but only indirectly through the exchange into dollars.

Thus, even for pound holdings, those variables that determine the choice between dollars and gold will also influence the amount of pounds a central bank is willing to hold. Similarly this is true for the other foreign exchange holdings. Therefore, for a central bank the primary decision is concerned with the choice between gold and foreign exchange since all foreign exchange holdings are affected by those factors that determine the price of gold in terms of dollars.

This is not to deny that if a country holds pounds or other reserve currencies, additional variables may have an influence on the ratio of foreign exchange to total reserves.

In order to minimize the complications resulting from the existence of a second re-

serve currency, we have excluded the sterling area countries which hold most of the external pound balances. And for the countries investigated, whose foreign exchange holdings are mainly in dollars but may also contain pounds, we have tested whether any factors that determine the position of the United Kingdom and its relation to other countries had an influence on the ratio of foreign exchange to total reserves. The only U.K. variable that showed a significant influence was the amount of a country's exports to the U.K. relative to its total exports.¹ We said, "... if a country exports primarily to reserve-currency countries, it is more likely to accept a low gold content of its reserves."² By holding foreign exchange instead of gold, it can help the reserve-currency countries to avoid import restrictions which would be most harmful to its own foreign trade.

In conclusion, we can say that the empirical analysis of Part II was intended to deal with and did deal with the key question how central banks allocate their international reserves between foreign exchange holdings and gold. We recognize that foreign exchange holdings represent a conglomerate of many different kinds of reserve assets.³ Of course, if we had more information on the composition of foreign exchange holdings of the different countries, we would probably be able to refine our empirical analysis and make more realistic assumptions for the hypothetical calculations in Part III.

REFERENCE

- H. A. Hagemann, "Reserve Policies of Central Banks and Their Implications for U.S. Balance of Payments Policy," *Amer. Econ. Rev.*, Mar. 1969, 59, 62-77.

¹ H. A. Hagemann, Table 2, p. 70.

² Hagemann, p. 65.

³ Even dollar holdings are not homogeneous, but contain a great variety of reserve assets with different interest rates, maturities and guarantees.

Maximum Bias in Social Saving Estimates Using Prices

By LLOYD J. MERCER*

The concept of social saving (the value of output produced by resources freed from other uses by an innovation) was introduced into economic history by Robert W. Fogel and has quickly become an important tool in empirical research (see Albert Fishlow, G. A. Gunderson). The meaning of social saving is best illustrated by consideration of a specific innovation, e.g., the railroad. Introduction of the railroad into the transportation network means (if the railroad is more productive than the next best alternative in the provision of transportation services) that fewer resources than before will be required to move a given bundle of goods and passengers between a given set of shipping points and destinations. Thus the following question is posed by attempts to measure social saving due to the railroad. What is the real resource saving as a result of being able to use the railroad rather than the next best alternative? Or alternatively, what is the gain in total real output using those resources freed by the railroad in their next best use?

The value in their next best use of those resources devoted to transportation is represented by the marginal cost of providing transportation services. The real resource saving of the railroad in providing freight transportation in any given year is:¹

$$(1) \quad (MC_a - MC_r)R$$

where

MC_a = the marginal cost of freight service by alternatives.

* The author is assistant professor of economics at the University of California, Santa Barbara.

¹ This assumes that the same quantity of transportation services would be demanded with and without the railroad. The economy is not allowed to adjust in the absence of the railroad to a less efficient transportation network. If this adjustment were made, it would result in a smaller quantity of transportation services demanded in the absence of the railroad. The measure of real resource saving suggested contains an upward bias for this reason.

MC_r = the marginal cost of freight service by railroads.

R = total ton-miles of the railroads.

The general formula for railroad social saving in passenger transportation can be written like (1) with the appropriate variables. Total railroad social saving would then be the sum of the freight and passenger components.

If estimates of social saving due to the railroad innovation are constructed on the basis of prices rather than marginal costs, the result is an estimate of total direct financial saving to those using the railroad which is not necessarily equal to the real resource saving (real social saving) due to the railroad (innovation).² The result of the use of prices rather than marginal costs to estimate social saving may be a significant downward bias in the social saving attributed to the railroad innovation.

The purpose of this note is to suggest a measure of the maximum bias resulting from the use of prices rather than marginal cost in the estimation of social saving. Knowledge of the magnitude of the possible bias is necessary for interpretation of the meaning of existing social saving estimates. As an example, an estimate will be made of the maximum bias from this source in the estimates by Fogel and Fishlow of railroad social saving in the United States.³ Recognizing the existence of the bias, but not measuring its magnitude, Peter McClelland (p. 123) concluded with respect to Fogel and Fishlow's

² When the differential between prices is equal in magnitude to the differential between marginal costs, the magnitude of direct financial saving will equal that of real social saving. When this condition is not met, direct financial saving (social saving estimated using prices) will not equal real social saving.

³ The Fogel and Fishlow estimates include only the direct saving from lower cost transportation of goods and passengers. They do not include any additional real resource saving (increased production) which may have resulted from the forward and backward linkage effects of the railroad.

results that, "Their central question, however—the net benefit to the economy from the existence of the railroad in 1859 or 1890—remains essentially as they found it: an unsolved mystery."

Both Fogel and Fishlow in estimating real social saving for freight shipments by rail (primarily) use the difference between the per unit price of shipment by alternatives and rail. Fishlow (pp. 90–92) estimates the direct social saving of the railroad for passengers in the same way.⁴ Because of this specification of the measure of per unit real social saving McClelland (p. 114) suggests that Fogel and Fishlow are "... *virtually assured of a small result.*" A small result may, of course, grossly understate the net contribution of the railroads to American economic growth, but the statement quoted starts one speculating on what exactly a *small* result might be.

To interpret the Fogel and Fishlow estimates of railroad social saving, we need to know the maximum bias resulting from their use of prices rather than marginal costs. This will allow the establishment of an upper bound for the magnitude of railroad social saving using their methods with marginal costs rather than prices. Bias is introduced into the Fogel and Fishlow measures of railroad social saving for freight transportation if (2) and (3) are not identical.

$$(2) \quad (P_a - P_r)R$$

$$(3) \quad (MC_a - MC_r)R$$

where

P_r = The price per ton-mile charged by railways.

P_a = The price per ton-mile charged by the alternative carrier (wagon or water transportation companies).

MC_r = Marginal cost of the alternative carriers for freight.

R = Total ton-miles carried by railways.

It is possible for a measurement using equation (2) rather than (3) to produce a small result only if,

⁴ Fogel does not estimate railroad social saving resulting from passenger use of the railroad rather than alternatives.

$$(4) \quad (P_a - P_r) < (MC_a - MC_r)$$

This requires,

$$(5) \quad (P_a - MC_a) < (P_r - MC_r)$$

which will result if railroads have a greater degree of monopoly power than alternative modes of transportation. Suppose that $P_a = MC_a$. Then the upper bound of divergence per unit shipped by rail between real social saving and the Fogel and Fishlow type measurement will be $P_r - MC_r$.⁵ However, with profit maximization $MC_r = MR_r$ for the quantity of freight services provided where MR_r = marginal revenue of the railways. It is a well-known proposition in microeconomics that $MR = P - P/\eta$ where η = the absolute value of price elasticity of demand. Thus,

$$(6) \quad MR_r = P_r - \frac{P_r}{\eta_r}$$

where η_r = the absolute value of the elasticity of demand for railroad freight transportation, and

$$(7) \quad \frac{P_r}{\eta_r} = P_r - MR_r$$

$$(8) \quad \left(\frac{P_r}{\eta_r} \right) (R) = D_f$$

where D_f = a measure of the maximum divergence between real railroad social saving attributable to freight transportation using the Fogel-Fishlow estimates of social saving with prices and the same estimates with marginal costs.

If $MC_r = 0$ and railroads maximize profit, we see from equation (6) that the absolute value of $\eta_r = 1$. If $MC_r > 0$ it is clear that the absolute value of η_r is larger and D_f is smaller.

The same argument can be made for railroad social saving attributable to passenger transportation. The measure of the maximum divergence between railroad social saving estimates using prices and those using marginal costs would be $D_p = (P_{rp}/\eta_{rp}) (P)$

⁵ Examination of (5) shows that if $P_a > MC_a$ the difference between (2) and (3) is reduced compared to the case where $P_a = MC_a$.

where D_p is the measure of the maximum divergence, P_{rp} is the passenger-mile price of rail travel, η_{rp} is the absolute value of demand elasticity for rail passenger travel and P is total rail passenger miles.

Let D_i be the measure of the maximum divergence between total railroad social saving estimated with prices and estimated with marginal costs, i.e., $D_i = D_f + D_p$. This will be a measure of the maximum divergence because D_i falls toward zero as the absolute value of demand elasticity in each segment of the market increases from the lower limit of 1. Where prices and marginal costs of alternatives are equal and elasticities of demand for the quantity of freight and passenger services produced by railroads are both 1, the components of D_i are easily obtainable. Total freight receipts of the nation's railways are D_f , i.e., $D_f = (P_f)(R)$. Passenger revenues are D_p , i.e., $D_p = (P_{rp})(P)$.

Fishlow (pp. 336-37) estimates total freight receipts as \$66.5 million and total passenger receipts as \$45.8 million in 1859. Thus, D_i is \$112.3 million in 1859. The U.S. Bureau of Census (p. 634) reports total freight and passenger receipts in 1890 (for the year ending June 30) as \$759.2 million and \$314.9 million respectively. Thus, D_i is about \$1,074.1 million in 1890.

Gross National Product (*GNP*) in 1890 was about \$12 billion (see Fogel p. 41). Hence, the upper bound of divergence between the Fogel estimate of 1890 social saving using prices (which neglects passenger traffic) and the result using marginal costs is about 9 percent of *GNP*. This, added to Fogel's (p. 219) highest upper bound estimate (7.1 percent), produces an upper bound of 16.1 percent of *GNP* for railroad social saving in 1890.

In 1859 *GNP* has been estimated as \$4.17 billion (see Fishlow p. 52). Thus, the maximum divergence between the Fishlow estimate of 1859 social saving using prices and the result using marginal cost is 2.7 percent of *GNP*. This added to Fishlow's upper limit estimate of 1859 social savings of about 5.4 percent (p. 93) results in an upper bound of 8.1 percent of *GNP* for railroad social saving in 1859.

The suggested measure of the maximum divergence of the Fogel and Fishlow price determined railroad social saving estimates from marginal cost based estimates makes Fogel's estimate about 44 percent of the maximum social saving and Fishlow's about 67 percent of the maximum. While these are relatively large divergences from the maximum social saving, the divergence from the actual is (probably) considerably less since marginal costs were in fact greater than zero.

How significant are social saving of 8.1 and 16.1 percent and how do these compare with the Fogel and Fishlow estimates? The following example is presented as one means of placing these numbers in perspective.⁶ If *GNP* in 1890 in the absence of the railroad had been 16.1 percent less than it actually was and the annual real growth rate was 3 percent, there would have been a lag of almost six years in achieving the actual 1890 *GNP* of \$12 billion.⁷ This compares with a lag of about two and one-half years if Fogel's highest estimate of social saving is used. If *GNP* in 1859 in the absence of the railroad had been 8.1 percent less than it actually was and the annual growth rate was 3 percent, there would have been a lag of slightly under three years in achieving the actual 1859 *GNP* of \$4.17 billion. This compares with a lag of almost two years if Fishlow's upper limit estimate of social saving is used.

The range of values established here shows that the Fogel and Fishlow railroad social

⁶ It should be remembered that the social saving estimates include only direct saving. While the author does not believe that additional saving generated by forward and backward linkages would significantly change the estimates of social savings, some readers may judge the reverse to be true. The example may be adjusted accordingly.

⁷ The *GNP* growth rate used allows *GNP* per capita to increase only about 1 percent or less per year since population grew 1.9 percent per year between 1890 and 1900 and 2.4 percent between 1860 and 1870 (see Elizabeth Gilboy and Edgar Hoover p. 258). The average rate of growth of per capita real product for the period 1800-1835 (as well as 1790-1860) is estimated at 1.3 percent (see Paul David p. 194). Thus the rate of growth used in the example is quite modest. Given the social saving rates, a higher growth rate would reduce the suggested lag in income growth in the absence of the railroad.

saving estimates do provide us with a measure of the contribution of the railroad to American economic growth. The benefit to the economy from the existence of the railroad in 1859 or 1890 is not an unsolved mystery. In terms of the question of "indispensability," any one person's evaluation of that contribution is not likely to be changed regardless of where the "true" contribution lies within the range established above. In a similar manner a range of values may be estimated for other innovations where prices rather than marginal costs were used to estimate social saving due to the innovation.

REFERENCES

- P. A. David, "The Growth of Real Product in the United States Before 1840: New Evidence, Controlled Conjectures," *J. Econ. Hist.*, June 1967, 27, 151-95.
- A. Fishlow, *American Railroads and the Transformation of the Antebellum Economy*, Cambridge 1965.
- R. W. Fogel, *Railroads and American Economic Growth*, Baltimore 1964.
- E. W. Gilboy and E. M. Hoover, "Population and Immigration," in Seymour E. Harris, ed., *American Economic History*, New York 1961.
- G. A. Gunderson, "Social Saving of Steamships," unpublished doctoral dissertation, Univ. Washington 1967.
- P. D. McClelland, "Railroads, American Growth, and the New Economic History: A Critique," *J. Econ. Hist.*, Mar. 1968, 28, 102-23.
- U.S. Bureau of the Census, *Eleventh Census of the United States, 1890*, Vol. 14, Part 1, Washington 1895.

The Structure of the Money-Expenditures Relationship: Comment

By DWIGHT M. JAFFEE*

In a recent article in this *Review*, Sam Peltzman has provided econometric estimates of the structural determinants of the velocity of money. In his formulation the velocity of money depends on a constant, a time trend, last period's velocity, and a lagged measure of real expenditures. The most notable results are that velocity does not depend on interest rates and that the implied income elasticity of demand for money is at least 1.8.

I shall develop two main points in this note. In Section I it is shown, using Peltzman's own behavioral hypothesis, that a key variable has been omitted from the estimated equations. When this variable is correctly included, the estimated coefficients change significantly and a fortiori the conclusions are radically changed. In Section II several questions are raised concerning the method of estimation and the interpretation of the time trend. The discussion is directed at a reconciliation of Peltzman's findings with previous empirical work on the determinants of velocity.

I. The Peltzman Model

The behavioral structure of Peltzman's model can be written as:¹

$$(1) \quad (E - M)^* =$$

$$a_0 + a_1T + a_2(E - P)_{-1} + a_3i_{-1}$$

$$(2) \quad \Delta(E - M) =$$

$$\gamma[(E - M)^* - (E - M)_{-1}]$$

* The author is assistant professor of economics at Princeton University. He wishes to thank his colleagues Lester V. Chandler and Stephen M. Goldfeld for commenting on a first draft of this note.

¹ We have retained Peltzman's notation wherever feasible. Equations of the specific form of (1) and (2) do not explicitly appear in the original because we have adopted at the outset the set of assumptions which Peltzman includes only as he proceeds in the develop-

E = logarithm of expenditures in current dollars

M = logarithm of money stock

P = logarithm of price deflator for E

T = linear time trend

i = logarithm of the interest rate

* = desired value

The time unit is a quarter

The model is thus seen to be in standard partial adjustment form. Equation (1) defines the desired value for velocity as a function of the indicated variables, and equation (2) specifies the actual change as proportional to the difference between the desired velocity and the actual velocity of the preceding period.²

Taking this model as given, we can now easily proceed to derive its testable form by substituting (1) into (2):

$$\begin{aligned} \Delta(E - M) &= \gamma a_0 + \gamma a_1 T \\ (3) \quad &- \gamma a_2 (E - P)_{-1} \\ &+ \gamma a_3 i_{-1} - \gamma (E - M)_{-1} \end{aligned}$$

Equation (3) is easily estimated pending only the specification of the empirical counterparts of the variables.³ Peltzman, how-

ment of the model. These assumptions relate to the logarithmic form of variables, the choice of the time unit, and the lagged timing of variables.

² Peltzman initially refers to the dependent variable in our equation (2) as the "intended" change in velocity, and denotes this by placing an asterisk superscript on the variable. This distinction between the actual change and the "intended" change is not made in his empirical tests, however, and thus our equation (2) specifies the actual change.

³ It should be emphasized that Peltzman's model assumes partial adjustment of velocity rather than partial adjustment of money demand with income given. The two assumptions lead to different testable equations as can be seen as follows. For the money demand adjustment model, we assume $\Delta M = \gamma(M^* - M_{-1})$ and then using equation (1) normalized on M^* as the desired

TABLE 1—REGRESSION RESULTS, 1953-I THROUGH 1967-II*

Money defined as:	Constant ($\times 100$)	$\Delta(E-P)_{-1}$	$\Delta(E-M)_{-1}$	Δi_{-1}	R^2	S_e	D.W.
A. Currency plus demand deposits = M_1							
A.1 $\Delta^2(E) =$.790 (5.8)	-.296 (-2.3)	-.649 (-4.8)	.007 (1.4)	.537	.622	2.04
A.2 $\Delta^2(E-M) =$.777 (4.9)	-.038 (-.26)	-.967 (-6.2)	.012 (2.2)	.521	.716	2.17
B. M_1 plus commercial bank time deposits = M_2							
B.1 $\Delta^2(E) =$.428 (2.9)	-.416 (-3.3)	-.406 (-3.8)	.009 (1.8)	.480	.659	2.13
B.2 $\Delta^2(E-M) =$.157 (.79)	-.135 (-.78)	-.717 (-4.9)	.015 (2.1)	.405	.894	2.10
C. M_2 plus time deposits at mutual savings banks and savings and loan associations = M_3							
C.1 $\Delta^2(E) =$.173 (.99)	-.320 (-2.3)	-.522 (-4.0)	.009 (1.8)	.490	.652	2.13
C.2 $\Delta^2(E-M) =$	-.159 (-.75)	-.063 (-.37)	-.781 (-4.9)	.013 (2.1)	.428	.798	2.08

* The equations were estimated using ordinary least squares; t -statistics are shown in parentheses and the R^2 is not corrected for degrees of freedom. Data sources are given in Peltzman (Table 1). Data revisions available April 1969 were used.

ever, goes two steps further. Because the variables still contain trend components, he takes first differences of (3) to obtain:

$$(4) \quad \begin{aligned} \Delta^2(E - M) = & \gamma a_1 - \gamma a_2 \Delta(E - P)_{-1} \\ & + \gamma a_3 \Delta i_{-1} \\ & - \gamma \Delta(E - M)_{-1} \end{aligned}$$

He then breaks the dependent variable into two parts, $\Delta^2 E$ and $\Delta^2 M$, and to "simplify" assumes $\Delta^2 M$ is zero, so that his specification of (4) contains only $\Delta^2 E$ as the dependent variable. This assumption is both unneces-

money demand function, we obtain upon arrangement of terms

$$\begin{aligned} \Delta(E - M) = & \gamma a_0 + \gamma a_1 T - \gamma a_2 (E - P)_{-1} + \gamma a_3 i_{-1} \\ & - \gamma (E - M)_{-1} + (1 - \gamma) \Delta E \end{aligned}$$

The velocity adjustment assumption, in contrast, leads to equation (3) which does not have the term in ΔE .

sary and unwarranted. It is unnecessary because the data for the money stock is of course available. It is unwarranted because omitting this variable will tend to bias the coefficients; more specifically, suppressing $\Delta^2 M$ will impart a downward bias to the interest rate coefficient.⁴

Table 1 shows the estimated coefficients for several equations which allow a comparison of Peltzman's original specification with the corrected formulation developed here. In all cases we show only the results using consumption expenditures as the measure of E (as this is Peltzman's favored specification) and the Treasury bill rate as

⁴ The direction of this bias is easily seen by moving the term $-\Delta^2 M$ to the right side of equation (4). Since the money stock and interest rate can be expected to be negatively correlated, the interest rate coefficient will be smaller when the money stock variable is suppressed.

the measure of i . Qualitatively the results are unchanged if we substitute net national product for E and/or the long-term government bond rate for i .⁵ Equations (A.1), (B.1), and (C.1) are estimates of equation (4) for three different definitions of money using Peltzman's assumption that $\Delta^2 M$ is zero.⁶ Equations (A.2), (B.2), and (C.2) are the corresponding estimates of equation (4) with the complete specification as given. The following generalizations can be made concerning the effect of properly including $\Delta^2 M$ in the dependent variable: in each case, the coefficient and t -statistic for the interest rate increased; the speed of adjustment (the absolute value of the coefficient of $(E-M)_{-1}$) increased and became more significant; the coefficient of real expenditures became smaller and insignificant; and the constant became smaller and, except for the narrow definition of money, insignificant.

It is apparent that the structural determinants of desired velocity (equation (1)) implied by our estimates stand in strong contrast with Peltzman's findings in three ways: (i) in line with almost all other studies of the demand for money, we find that the interest rate is a significant determinant of velocity, although the long-run elasticity is relatively small; (ii) there is no indication of an effect from real expenditures over and above the unitary elasticity assumed in the velocity specification; (iii) the time trend is a significant determinant of velocity only for the narrowest definition of money.

⁵ Peltzman's choice of the expenditure variable rests on considerations related to the correct proxy for permanent income. In the context of a partial adjustment model this is unnecessary, however, because the model itself implies a lagged adjustment to the expenditure variable, and the distribution of the lag may be interpreted as permanent income weights. This also suggests that the Koyck lag distribution implicit in the model may be overly restrictive.

⁶ These results differ somewhat from the corresponding estimates shown by Peltzman because we have used data incorporating revisions for both the expenditure and money stock data that were not available earlier. It is surprising that the coefficients are not more robust to the data revisions, and in Section II this is attributed at least in part to the choice of second differences for the specification.

II. *The Significance of Time Trends*

Although our estimates have shown the interest rate to be a significant determinant of velocity, there still remains the need to reconcile the relatively low long-run interest elasticity and the significance of the time trend when money is narrowly defined with previous empirical studies of the demand for money. While one could quarrel with the details of the specification of variable lags and the like, these factors are probably not critical to the long-run properties of the equation. More likely explanations lie in Peltzman's method of estimation and his interpretation of the time trend.

The source of the statistical problem concerns Peltzman's use of second differences in order to avoid multicollinearity in an equation with trend-like variables. Even in general practice this solution may be of limited value since what one gains in excluding multicollinearity is likely to be offset by deterioration in the quality of the data as one uses changes over short periods of time.⁷ For the model at hand this case for second differences is even weaker because the specification in first difference form already includes a time trend, and it has been shown (see Tinter, pp. 301-304) that this effectively removes the trend from the other independent variable. Furthermore, at least in principle, the choice between first and second differences is properly based on the properties of stochastic error term. This is particularly important in estimating partial adjustment models because serial correlation of the error term together with a lagged dependent variable may be the source of inconsistent coefficient estimates.⁸

As a direct test of the sensitivity of the

⁷ We have already noted the coefficients show some change when estimated on revised data, and this may be a symptom of using second differences.

⁸ Furthermore, the serial correlation of the error may cause simultaneous equations bias even though the interest rate is lagged one period in the model. This suggests using generalized least squares and possibly two-stage least squares to obtain more reliable estimates. See Franco Modigliani et al. for further discussion of this point.

results to the use of second differences, we have estimated the equivalent of equation (A.2) in first difference form:

$$\begin{aligned} \Delta(E - M) &= .726 + .414T \\ &\quad (2.8) \quad (4.1) \\ (5) \quad &-.109(E - P)_{-1} - .407(E - M)_{-1} \\ &\quad (-2.4) \quad (-4.3) \\ &+.014i_{-1} \\ &\quad (3.7) \\ R^2 &= .283 \quad S_e = .642 \quad D.W. = 1.70 \end{aligned}$$

The impact elasticity of the interest rate remains essentially unchanged, although it becomes still more significant, but the speed of adjustment falls by more than half. This implies that the long-run interest elasticity more than doubles when first differences are used. Furthermore the standard error of estimate is about 10 percent smaller in this case.

Turning finally to the time trend, Peltzman interprets this variable as a proxy for noninterest rate factors, such as Milton Friedman's economic stability, which may have caused the postwar increase in velocity. A more specific hypothesis, and one sup-

ported by the fact that the time trend enters significantly only with money narrowly defined, is that the trend is measuring the secular shift from demand deposits to time deposits. This in turn may be explained by factors such as the rising time deposit rate, the introduction of Certificates of Deposit, and the increasing use of credit cards. Indeed, work for the *FRB-MIT-PENN* model by Modigliani, Rasche, and Cooper indicates that the time deposit rate is as important a determinant of money demand in the long run as the Treasury bill rate. Or, in other words, in this interpretation the time trend may simply be a proxy for additional interest rate effects and these must be properly taken into account in calculating the true interest elasticity.

REFERENCES

- F. Modigliani, R. Rasche, and P. Cooper, "Central Bank Policy, Interest Rates, and the Money Supply," *J. Money, Credit and Banking*, forthcoming.
- S. Peltzman, "The Structure of the Money-Expenditures Relationship," *Amer. Econ. Rev.*, Mar. 1969, 59, 129-37.
- G. Tinter, *Econometrics*. New York 1965.

The Structure of the Money-Expenditures Relationship: Reply

By SAM PELTZMAN*

Dwight Jaffee's most important criticism of my article is that the empirical estimates of the model gratuitously excluded a variable, and that this biased downward my estimate of the long-run interest elasticity of velocity. Since the most notable difference between my article and the bulk of the literature on the demand for money is my finding of an essentially zero interest-elasticity of velocity, Jaffee's criticism is surely germane. However, Jaffee has misinterpreted my model, and, even if his misinterpretation is accepted, his empirical results do little violence to the findings.

Jaffee's misinterpretation of my model can be seen by comparing his equation (2) with equation (1) in my article, which I rewrite, adopting Jaffee's notation:

$$(1) \quad \begin{aligned} &[\Delta(E - M)]^* \\ &= \gamma[(E - M)^* - (E - M)_{-1}]. \end{aligned}$$

The dependent variable is the *intended*, not the actual, change in the logarithm of velocity. The difference between my (1) and Jaffee's (2) is subtle, but important. The distinction bears elaboration here, because Jaffee's misinterpretation convinces me that my original presentation was too succinct.

Equation (1) states that the representative money holder intends to change velocity by some fraction of the difference between desired and actual velocity. This intended change can be thought of as having two components: 1) The money holder expects some change in his money holdings, and 2) given this, he plans a change in expenditures sufficient to accomplish the intended change in velocity. For example, a money holder who wishes to increase velocity by 5 percent and expects his money holdings to rise by 2 per-

cent will plan to increase expenditures by 7 percent. Now, it will be recalled that my article focuses on the effect of money on expenditures, so I isolated the expenditures component of the intended change in velocity by writing (1) as (again in Jaffee's notation):

$$(2) \quad \begin{aligned} (\Delta E)^* &= \gamma[(E - M)^* - (E - M)_{-1}] \\ &+ (\Delta M)^*. \end{aligned}$$

In order to make (2) operational, some assumptions about $(\Delta E)^*$ and $(\Delta M)^*$ were required. I chose to set the intended expenditure change, the variable of primary interest, equal to the actual change and to set the expected change in money equal to a constant. Jaffee would have me set the latter change equal to its actual value as well. This is unnecessarily restrictive. For the aggregate of money holders, the actual change in money is, after all, largely determined by the monetary authorities. Unless spending units forecast exactly what Federal Reserve monetary policy is going to be over the period for which they are making expenditure plans, the assumption of equality between expected and actual money changes may not be useful. My procedure can surely be improved upon: for example, one might set $(\Delta M)^*$ equal to an empirically determined weighted average of past money changes. But, to assume, as Jaffee has in effect done, that money holders have perfect foresight about Federal Reserve policy raises as many difficulties as it eliminates. For example, suppose money holders do not have such perfect foresight, and the Federal Reserve conducts a counter-cyclical monetary policy. This will introduce a positive bias into the coefficients of $\Delta(E - P)_{-1}$ in Jaffee's Table 1. To illustrate: let the Federal Reserve observe a low (high) value of $\Delta(E - P)$ today and react by accelerating (retarding)

* Visiting associate professor of business economics, Graduate School of Business, University of Chicago.

the money supply next quarter in an amount not entirely anticipated by money holders. Then $\Delta(E-P)_{-1}$ and $\Delta^2(E-M)$ will tend to move together, without this movement having been intended by the money holders whose behavior the model purports to describe. Indeed, this may explain why the coefficient of $\Delta(E-P)_{-1}$ increases algebraically when Jaffee goes from my model to his version of it.

More important than these theoretical points is the substance of Jaffee's empirical results. Accept, for these purposes, Jaffee's version of my model. What important difference does this really make? Jaffee does indeed report significantly positive interest rate-elasticities of desired velocity. However, note their magnitudes. For all money supply definitions, and in first or second difference form, the estimated interest-elasticity of desired velocity never exceeds .035. I reported a maximum value of .014 (p. 132, fn. 5), in my 1969 article. In this context the difference between .035 and .014 or zero is negligible. Specifically, an elasticity of .035 cannot account for any substantial part of the postwar rise in velocity. Further, Jaffee's estimates for conventional money retain the significant time trend in desired velocity in both first and second difference form. It was precisely my point "...that those who have explained the postwar rise in velocity by interest rate movements have confounded interest rates with some other trend-related phenomena" (p. 136). No substantially different conclusion

can be drawn from Jaffee's own results.

I agree with Jaffee's interpretation of the time-trend coefficients for various money definitions, namely that there has been a secular shift from demand to time deposits. I made just that point (p. 137, fn. 16). However, one must doubt that rising time deposit interest rates can explain a substantial part of this movement. Time deposit rates have, after all, moved up with other interest rates, and both Jaffee's estimates and mine show interest rates to be of negligible importance in explaining the postwar rise in velocity. Something else, perhaps improvements in the technology of banking or improved expectations about economic stability, must be responsible for this shift from demand to time deposits.¹ The shift to time deposits might then simply be part of an overall shift from money to non-money forms of wealth. In any case, since Jaffee's results no less than mine appear to rule out an interest rate explanation for the postwar rise in velocity, the task for future research remains to identify empirically the trend-related source of the postwar rise in velocity.

REFERENCE

- S. Peltzman, "The Structure of the Money-Expenditures Relationship," *Amer. Econ. Rev.*, Mar. 1969, 59, 129-37.

¹ It is often overlooked that demand deposits usually pay interest, albeit in non-explicit forms. The essentially zero interest-elasticity of velocity may be due to a sufficiently rapid adjustment of demand deposit interest rates to other money market rates.

An Econometric Model of Development: Comment

By SARA S. BERRY*

Sixty years ago, Joseph Schumpeter began his *Theory of Economic Development* with a discussion of the partial character of all economic analysis. "Economic development is . . . simply the object of economic history which in turn is merely a part of universal history . . . Because of this fundamental dependence of the economic aspect of things on everything else, it is not possible to explain *economic* change by previous *economic* conditions alone" (p. 58). Schumpeter was writing primarily about the historical experience of European societies who shared many common cultural and institutional features; his point is, if anything, even more relevant for contemporary students of economic development who try to explain patterns of economic change for a much greater variety of societies. Yet, although many economists have recognized that differences in the economic performance of different societies cannot be explained by economic factors alone, few have attempted systematic analyses of the respective roles of economic and non-economic factors in generating economic change. Consequently, the literature on economic development offers few general propositions about the quantitative importance of political and cultural factors for development, or about the way in which such factors interact with economic forces to alter the structure or volume of an economy's productive capacity.

In their article "An Econometric Model of Socio-Economic and Political Change in Underdeveloped Countries," Irma Adelman and Cynthia Taft Morris have undertaken an important step toward filling this crucial gap in the literature. Specifically, they have attempted, using stepwise regression analysis, to estimate the importance of a large number of variables for explaining differences in the economic performance of sev-

enty-three non-Communist countries during the 1950's and early 1960's.¹ They then used the regression coefficients to calculate multipliers measuring the impact of changes (of one standard deviation) in eighteen of the independent variables on a country's "development potential."

They included a number of social and institutional variables which are not readily quantifiable and explored relationships about which there are few widely accepted theoretical propositions. The significance of their conclusions about how and why countries develop economically depends not only on the form of their statistical analysis, but also on their identification and quantification of relevant variables and on their interpretation of the statistical associations discovered. In all these respects the authors faced formidable problems, which they have not always solved very successfully.

I. The Independent Variables

In constructing quantitative indicators for various institutional and social factors which may affect countries' economic development, Adelman and Morris follow the general procedure of classifying countries into several categories according to the factor in question and then assigning values to the categories at arbitrary intervals on a linear scale.² They justify this procedure on the grounds that all their variables are continuous and that "the only difference between our case and the more usual one is that deficiencies in information led us to scale our variables by means of a yardstick which has demarcations at coarse intervals only (e.g.,

¹ In their article, Adelman and Morris label their original dependent variable "development potential" but it consists of a classification of countries according to several indicators of past economic performance. (p. 1188, fn. 6)

² Detailed explanations of how Adelman and Morris defined and measured the variables in their model are given in their book, *Society, Politics, and Development*.

* The author is an assistant professor of economics at Indiana University.

feet or yards instead of inches)" (p. 1204). This argument is misleading since, for a number of variables, the "intervals" to which Adelman and Morris have assigned equal numerical values do not represent equal quantitative changes in the underlying variables. For example, in constructing their index of "social mobility" (X_{25}), the authors assigned scores of 90 to countries with more than 40 percent of the population aged 5 to 19 in school and more than 10 percent of the active males in "middle class" occupations; 50 to countries with 25 to 40 percent of the 5-19 age group in school and 5-10 percent of the active males in middle-class occupations; and 10 to countries not classed in the above groups (pp. 34-35). Thus, the 40 point difference between the first two groups represents a 1 to 15 percent difference in the proportion of school-aged children in school, while the 40 point difference between the latter two groups represents a 1 to 25 percent difference in this indicator. In addition, some countries received lower scores than these criteria warranted because they were judged to have "prohibitive cultural and ethnic barriers to upward social mobility." This criterion further obscures the meaning of an interval on the social mobility scale because it is not operationally defined. In short, what the intervals seem to measure is as much the authors' opinion (or the opinions of anonymous "country experts") about the extent of social mobility in different countries as any objective evidence on the subject.¹

Some of the independent variables included in Adelman and Morris's model are not only arbitrarily measured but also ambiguously defined. A case in point is their treatment of dualism, an overworked term

in development literature which has long ceased to have any precise meaning. The original proponent of the term, J. H. Boeke, defined it as "the clashing of an imported social system with an indigenous social system of another style" (p. 4). Others have used the term to characterize economies with various kinds of sectoral differences, e.g., in technology and factor proportions, in the use of money and exchange, in the institutions available for capital accumulation and technical change, or in people's attitudes toward work, profit and efficiency. In order to use the concept of dualism in a quantitative model, one is faced with the task of sorting out various possible definitions into observable variables whose relationship to development potential can be interpreted meaningfully.

Adelman and Morris's definition does not accomplish this task, but tends to reflect the confusion in the literature. For example, they assign the lowest score on their dualism scale to the "largely agrarian societies having extremely small exchange sectors."⁴ At the other end of the scale are "countries with continuous interaction between modern and non-modern elements" (p. 1214). Apparently we are to assume that the larger the exchange sector and the less agrarian the society, the greater the degree of interaction between modern and non-modern elements. But it is not clear, either in theory or in practice, that the exchange sector always grows at the expense of the agrarian sector. In many West African countries, for example, commercial activity and the use of money are widespread, yet the majority of the population is still engaged in agriculture. Indeed, the authors appear to recognize this fact in discussing intermediate points along the scale: "countries in which the growth of an indigenous small-scale cash-crop sector using conventional techniques evolves at the expense of a traditional subsistence sector" rank relatively high on the scale. This sug-

¹ Moreover, by assigning arbitrary scores to different cultural or political characteristics of a society, the authors interject their own preferences into the estimated relationships between those characteristics and the society's economic performance. Another observer who chose to "evaluate" the three levels of social mobility defined above at scores of 10, 5 and 1 would obtain a different estimate of the regression coefficient relating differences in social mobility to differences in the authors' index of "development potential." For example, see R. D. Luce and H. Raiffa's discussion of Arrow's paradox (pp. 336-37).

⁴ Adelman and Morris's terminology is somewhat confusing here—they state that "positive changes in dualism involve reductions in the cleavage between modern and traditional sectors" (p. 1194) or, the less dualistic a society, the higher the score it receives.

gests that it is the extent to which goods are sold for money, rather than the composition of output or the occupational structure which best reflects the degree of interaction between modern and non-modern elements. If this is what the authors have in mind, however, it is difficult to see why they argue that "societies in which a foreign-financed and directed modern sector is superimposed upon a predominantly agrarian society" should rank below countries where cash-cropping is widespread among the indigenous population. In countries such as Zambia or Malawi (which rank low on Adelman and Morris's scale), many of the indigenous population regularly exchange goods and services (e.g., labor services) with foreign-financed and directed firms while continuing to engage in traditional agriculture. Thus there is a good deal of monetary exchange between modern and non-modern elements. Evidently, the authors had some other form of interaction in mind, but they do not state it explicitly. They have combined so many criteria (not all of which are closely intercorrelated) in their definition of dualism that is not clear what they are trying to measure or what significance ought to be attached to their statistical estimates of the associations between dualism and other variables in the model.

Moreover, in ranking individual countries on their dualism scale, Adelman and Morris have not applied their own criteria very consistently. For example, South Africa ranks well above Zambia, although in both economies a "foreign-financed and directed modern sector" exists side by side with a low-income agrarian sector; Uganda, where the bulk of the cash crops are produced on small farms by indigenous farmers, is scored below Rhodesia and Kenya, both of which have substantial foreign-financed and directed sectors. Since the authors do not justify their rankings of individual countries in detail, it is impossible to say what other criteria they used. However, available data suggest that South Africa does have a higher per capita income than Rhodesia which, in turn, has a higher per capita income than does Uganda. If, in effect, Adelman and

Morris's index of dualism measures levels of economic performance rather than some independent structural characteristic of the economies in question, then the strong correlation they find between dualism and development potential is neither surprising nor very instructive.

I have discussed the ambiguities in Adelman and Morris's definition and measurement of dualism at some length to illustrate what seems to be a fairly extensive problem in their analysis. There is not room in the present discussion to examine each of the nineteen variables in their model, but dualism is not the only one whose meaning is ambiguous because the authors have either combined several types of information in one index or defined variables in such a way as to assume relationships they are trying to measure.⁵

II. *The Regression Coefficients*

Even for those variables which are unambiguously defined, it is not always clear why the authors interpret their results as they do. In reading their initial premise, that "there are no firmly validated theories of the process of socio-economic and political change [so that] we consciously avoided a priori specification of the functions we wished to fit" (p. 1184), one anticipates a fairly thorough discussion of alternative interpretations of these functions once they are fitted. However, Adelman and Morris's explanations of their regression coefficients and the multipliers derived from them are very brief. In many cases they consider only one of several possible interpretations of a statistically significant relationship. Take, for example, the first regression equation which the authors present (after the discriminant

⁵ In other instances they have simply not taken the trouble to express themselves clearly. In defining "the importance of the indigenous middle class," one presumes that the authors do not really mean that "this classification is based upon the relative size . . . of indigenous people in middle-class occupations" (p. 1214), but the fact that they say so does not increase the reader's confidence in their other stated opinions or conclusions.

function which is used to construct the dependent variable):

$$(1) \quad X_{13} = .60X_4 + .33X_{21} \quad R^2 = .57$$

where X_{13} is the degree of improvement in financial institutions; X_4 is the degree of industrialization; and X_{21} is the extent of dualism. According to Adelman and Morris this equation shows that increased demand for credit in both industry and agriculture has a favorable effect on the performance of financial institutions. Since, as we have seen, dualism is not quite the same thing as the agricultural sector's demand for credit, their interpretation of the second regression coefficient seems a bit arbitrary. More important, however, is their failure to explain why the line of causation should run from industrial expansion to improved financial institutions rather than vice versa. In general terms, it seems just as plausible to argue that the growth of the banking system facilitates industrial expansion by helping to mobilize potential savings and to inform investors of their availability. In fact, in their 1967 study, Adelman and Morris argued that changes in structural variables might be considered causes of changes in economic performance because "it is difficult to see how, in the short run at least, causality could run in the opposite direction" (p. 175). Evidently they have changed their minds; one would be interested in their reasons for doing so.

Another difficulty of interpretation arises out of the authors' technique of analysis.⁶ Stepwise regression is useful in this kind of study in that it permits one to select from a large number of variables the few which are most closely associated with a particular dependent variable. Moreover, the technique selects first that independent variable which explains the greatest proportion of the variation in the dependent variable; second, the independent variable which explains the largest share of the remaining unexplained variation in the dependent variable, and so on. The disadvantage of stepwise regression

is that in selecting variables for inclusion in the regression equation, if any of the independent variables are intercorrelated so that they affect the dependent variable jointly as well as independently, the technique attributes all of the joint effect to the variables selected for inclusion. Thus, some variables rejected by the stepwise regression analysis as not adding significantly to the R^2 for a particular relationship may, in fact, have an important influence on the dependent variable. In the present example, if a rejected variable (say, X_2 , the gross investment rate) is intercorrelated with X_4 , the regression coefficient for X_4 includes the joint effect of X_2 and X_4 on X_{13} . Thus, the conclusion that the gross investment rate does not have a significant effect on improving financial institutions in developing countries may be incorrect.

III. *The Multipliers*

In the last section of their paper, Adelman and Morris provide "a statistical explanation of the economic and noneconomic forces which directly and indirectly determine a country's capacity for economic growth" (p. 1184). Specifically, they use the regression coefficients of the model to calculate multipliers which estimate the quantitative impact of changes in different socio-economic and political variables on development (measured by the discriminant function), and from the multipliers they draw conclusions about the relative importance of different variables for understanding how and why economies develop. Their conclusions seem questionable on several grounds. In the first place, there is no more evidence for the multipliers than there was for the regression coefficients concerning causal sequences. One may argue just as plausibly from their data that development will generate significant changes in outlook, economic structure and financial institutions as vice versa. Second, because their regression coefficients may include joint effects of variables not included in the equations (to say nothing of those not specified to begin with), the meaning of the multipliers is somewhat ambiguous. For example, the multiplier for the improvement

⁶ I am indebted for this point to my colleague Jeffery Green.

in financial institutions, 2.309, does not mean that if the rate of improvement in financial institutions rises by one standard deviation then development performance will rise by 2.309 standard deviations *ceteris paribus*, because other things are not necessarily equal. But if the multipliers cannot be interpreted in this fashion, it is not clear how they provide us with a valid ranking of the independent variables in terms of their impact on development.

Finally, the ambiguities in the authors' initial definitions and measurements of the independent variables are in no way mitigated by the statistical analysis. Since we still do not know, for instance, whose outlook weighs most heavily in Adelman and Morris's index of modernization of outlook or what kinds of information they used to measure differences in outlook, their conclusion that greater modernization of outlook will have a significant impact on economic development is not very enlightening. In the case of dualism, insofar as the authors' classification of countries was based on dif-

ferences in economic performance, their conclusion is tautological; otherwise it is merely confusing. In short, although Adelman and Morris have undertaken an important and formidable task, they do not seem to have solved enough of the problems inherent in any attempt to quantify interactions between economic and noneconomic forces to have produced significant or conclusive results.

REFERENCES

- I. Adelman and C. T. Morris, "An Econometric Model of Socio-Economic and Political Change in Underdeveloped Countries," *Amer. Econ. Rev.*, Dec. 1968, 58, 1184-1218.
- , *Society, Politics and Economic Development*. Baltimore 1967.
- J. H. Boeke, *Economics and Economic Policy of Dual Societies*. New York 1953.
- R. D. Luce and H. Raiffa, *Games and Decisions*. New York 1957.
- J. A. Schumpeter, *The Theory of Economic Development*. Cambridge, Mass. 1936.

An Econometric Model of Development: Comment

By PETER ECKSTEIN*

In the December 1968 issue of this *Review*, Irma Adelman and Cynthia Taft Morris present a cross-section analysis of the recent growth experience of underdeveloped countries. Their "econometric model of development" seeks to supplement economic analysis and country case studies as means of isolating those policies most likely to promote rapid development. The argument of this note is that their model not only fails to deal directly with some of the inherent problems of cross-country comparisons but also neglects some of the prime rules of scientific inquiry and offers, therefore, no reliable guide to policy-making. In commenting on their work, I shall also make reference to their book, *Society, Politics and Economic Development*, and to their journal article which appeared in May 1968.

The model has three stages: a discriminant analysis of 73 non-Communist underdeveloped countries grouped according to "development potential"; a series of multiple regression equations for the four explanatory variables that emerge as important from that analysis; and a set of "multipliers," based on these two analyses and designed to measure the impact on "development potential" of some 18 political, social, and economic variables. Each of these stages will be examined in turn.

I. The Discriminant Function

Discriminant analysis is a method of finding those characteristics which most clearly set off the members of one group from the members of other groups. Like regression analysis, it is a means of explaining or pre-

dicting a dependent variable (here called a criterion variable) in terms of several independent (or prediction) variables. Unlike regression analysis, only the prediction variables of discriminant analysis are defined continuously. The criterion variable is the group membership of the observation where this membership is known beforehand for all the cases in the sample.¹

Adelman and Morris begin their analysis by classifying the countries by their average annual rates of growth of per capita Gross National Product (*GNP*) between 1950-51 and 1963-64: "high" for rates above 2 percent, "intermediate" between 1 and 2 percent, and "low" below 1 percent. They explicitly reject, however, growth of *GNP* per capita as a measure of the "prospects for sustained economic growth" on the grounds that there are many cases of "growth without development." They therefore modify the membership of these groupings in several ways (discussed below) in order to arrive at a classification according to high, intermediate, and low "development potential." Although they begin with 29 potential prediction variables, their final discriminant function requires only four to assign each of the seventy-three countries a score closely approximating the mean for its own group and to account for "over 97 percent of the overall variance in membership." The four prediction variables are "the degree of improvement in financial institutions," "the degree of modernization of outlook," "the extent of leadership commitment to economic development," and "the degree of improvement in agricultural productivity."

Unfortunately, the impressive statistical power of the function seems largely the result of the ways in which the authors have defined and measured the criterion variable and the prediction variables.

* Assistant professor of economics, University of Michigan, and research associate, Center for Research on Economic Development. I am indebted to Elliot Berg, Lester Taylor, Wolfgang Stolper, and Shanti Tangri for careful and critical readings of earlier drafts. This note represents a condensation of my larger discussion paper, "Quantitative Measurement of Development Performance."

¹ See, for example, Kendall's *A Course in Multivariate Analysis*.

Defining "Development Potential"

There are two general reasons for regretting the decision to move from individual estimates of country growth rates to groupings of countries by development potential. By arbitrarily dividing the sample into only three groups, they have lost information about the substantial differences in growth performance within each of the groups.² Moreover, by attempting to transform groupings by growth rates into groupings by development potential, the authors have rested the entire analysis on a concept that could easily elude the most careful efforts at definition and measurement.

In practice, however, the authors neither define development potential prior to the analysis nor measure it independently of the prediction variables. They move from estimates of growth of per capita *GNP* to their concept of development potential in three distinct steps, each of which reduces the extent to which the variable being explained is observed independently of the variables being employed to explain it.

First, six countries with high growth rates are declared "unclassified" as to development potential and initially excluded from the analysis. Two—Panama and Burma—are completely unexplained, while four are African countries "in which, as of about 1960, over 90 percent of the population was in the traditional subsistence sector where per capita *GNP* had not changed significantly since 1950" (Adelman and Morris, 1967, pp. 89–90). For these four, then, poor performance on one of the prediction variables—improvement in agricultural productivity—was the explicit basis of the declassification, but five of the six countries are predominantly low on all four prediction variables.

Next, because they are interested in "development" rather than "growth," the authors seek to eliminate from the high-potential category those countries whose high

growth rates "had only a limited overall impact within the economy." They therefore relegate to the intermediate category eleven fast-growing countries that do not rank "at least moderately high" with respect to five out of seven "economic performance characteristics."³ Unfortunately, all seven of these characteristics are included among the twenty-nine variables which the computer later scans in order to build the discriminant function. Not surprisingly, two of the seven—improvement in financial institutions and improvement in agricultural productivity—that are here used to categorize countries as to development potential are also among the four that are eventually "found" to be the best predictors of the categories to which countries belong.⁴

The third step by which growth-rate groupings are transformed into development potential groupings is the most disturbing. On the basis of the rankings already described, the authors develop an initial discriminant function which contains four prediction variables—three of the four mentioned above, plus "degree of improvement in physical overhead capital" instead of the agricultural productivity variable. When scores on this discriminant function are calculated for the sixty-seven countries in the analysis, twelve countries are found to have ratings on the prediction variables more characteristic of countries in different categories of development potential, while three are marginal.

Adelman and Morris react to these discrepancies by observing that their confidence in the initial classifications for potential was "not especially high" in the first place (Adelman and Morris, May 1968, p. 277). They do not explain how their discriminant function could possess more validity than the criterion variable from which it is directly derived, or how their prediction variables—

² The seven are: change in degree of industrialization; improvement in agricultural productivity; improvement in physical overhead capital; improvement in financial institutions; improvement in the tax system; and improvement in human resources.

⁴ Examination of the ratings for the high-growth countries confirms that those rated low on either of these two characteristics were far more frequently demoted to the "intermediate potential" category.

³ For example, Turkey (with a per capita *GNP* growth rate of 2.1 percent) is grouped (and therefore assumed identical in "potential") with Greece (5.9 percent) but not with Costa Rica (1.9 percent).

apparently measured no more carefully or objectively than the criterion variable—could greatly exceed it in validity. Nevertheless, they proceed to employ both the discriminant function and the initial classifications of the prediction variables not only to classify the six countries for which development potential had been treated as unknown, but also to declassify the fifteen countries which the function did not fit well. They treat these new classifications and declassifications not as *hypotheses* but as *the equivalent of data*—that is, as the input to (or exclusions from) another discriminant analysis. The discriminant function resulting from this analysis is then used to reclassify the unknown development potential of the fifteen countries, and these classifications are treated as part of the data input to a third discriminant analysis.

When this procedure is complete, all seventy-three of the countries have been classified—although for twenty-one countries observed growth performance has been treated as irrelevant, and classification has been based *entirely* on ratings on the four prediction variables included in the first two discriminant functions. When the third and final function is calculated from those classifications, it accounts for “over 97 percent of the overall variance,” and it renders “the separation between groups . . . considerably better . . . and the dispersion within the groups . . . substantially reduced” (Adelman and Morris, May 1968, p. 278). This should not be surprising since the new function employs three of the four prediction variables used in classifying so many of the observations.⁵

Thus, there are three operations by which a classification of countries by growth rates

evolves into a classification by development potential, and each operation brings the implicit definition of high development potential closer to being nothing more than high performance on the four final prediction variables. There can, of course, be no scientific interest in even the strongest of statistical relationships between variables that have been defined to be similar.

Defining and Measuring the Prediction Variables

At least some of the statistical power of the discriminant function is based on underlying associations between the four prediction variables and the rate of growth of *GNP*.⁶ Unfortunately, we cannot conclude from these associations that any of the four is an independent cause of faster economic growth. Two of the variables could easily have been influenced by the same empirical phenomena that lay behind the growth-rate categorizations; the other two may merely serve as proxies for more basic explanatory variables.

“Leadership commitment to economic development” is one of several qualitative or “judgmental” variables; these were initially quantified by the authors on the basis of recent country studies and subsequently checked by country experts (largely officials of the U.S. Agency for International Development). Countries were ranked highest on leadership commitment if the government and central economic leadership had undertaken “concerted efforts” to promote growth, “serious attempts” to alter “institutional arrangements unfavorable to growth” and “some reasonably effective development planning” (Adelman and Morris, Dec. 1968 p. 1216; 1967, p. 80). The difficulty is that a judge must rely on indirect evidence to decide whether all that commotion over planning and development really represents “concerted” and “serious” effort, and the best indirect evidence may be whether or not

⁵ In all there are 32 reclassifications in which 26 of the 73 countries end up in new categories. In my discussion paper I have shown that the final discriminant function produces very similar scores for countries with high and intermediate rates of *GNP* growth, though the low-growth countries generally score lower. The ability to distinguish between countries ranked high and intermediate on development potential is largely achieved when the first discriminant function is used to reclassify the 6 originally unclassified countries and to declassify the 15 which fit it poorly.

⁶ The simple correlation coefficients with rate of growth are: .35 for leadership commitment, .61 for agricultural productivity, .54 for modernization, and .69 for financial institutions. The statistics are taken from Adelman and Morris (1967).

the effort is succeeding. Can a leadership really be judged "committed," can planning really be judged "effective," can institutional reform really be judged a contribution to growth, if the country is in fact not growing? Until more independent measures can be developed, we must suspect that unavoidable biases of judgment, not underlying causal relationships, account for some of the association between leadership commitment and growth.

A second prediction variable, "improvement in agricultural productivity," was intended to be a measure of total factor productivity. For economies dominated by agriculture, even accurate and complete data would be expected to show a high correlation between output per worker (or per factor) in agriculture and output per worker (or per capita) in the economy as a whole. We cannot find a great deal of theoretical interest in an empirical finding that across countries in which agriculture accounts for most of *GNP*, agricultural productivity and *GNP* growth rates are highly correlated.

In practice, the authors concede that even "reasonably reliable" data were often not available for either variable—and particularly not for the many low-income African countries—and they were forced to rely on purely "qualitative indications" or "estimates" (Adelman and Morris 1967, pp. 87–89, 104–108). There must have been available very little concrete evidence, then, on which to move from estimates of growth of output per worker in agriculture to quantitatively different estimates of the growth of output per factor in agriculture or of output per capita in the larger economy.

"Modernization of outlook," another judgmental variable, could easily be a proxy for such basic economic factors as level of *GNP* per capita or the extent of subsistence agriculture. In particular, a high negative correlation of "modernization" with the "size of the traditional agricultural sector" ($-.82$) implies very few differences in the rankings on the two variables. Those differences may have been because of actual differences in popular "outlook," but they also may have

been because the concomitants of high or low growth colored the interpretations of the experts. Indeed, the fact of growth itself could easily have been taken as evidence of popular support for "programs of political and economic modernization"—one of the major components of the definition of modernization of outlook. Even small differences of this sort would be sufficient to account for the greater explanatory power of this variable when it enters the discriminant function.

The variable "improvement in financial institutions" could easily be a surrogate for the "widespread pattern of dynamic improvements" that the authors find "well correlated" with it—"changes in the degree of industrialization, improvements in agricultural productivity, improvements in tax institutions, and to a lesser extent, improvements in human resources" (Adelman and Morris, May 1968, p. 269). Its statistical superiority over these other variables (several of which were also involved in the definition of the criterion variable) may be partly explained by the fact that one of its two components is "the approximate increase in the real value of private domestic liabilities to the banking system" (Adelman and Morris 1967, p. 121), apparently measured absolutely, not in relation to a growing *GNP*. Thus, even if all countries maintained a constant ratio of financial liabilities to *GNP*, this measure would be larger for the faster-growing economies.

Our examination of the four prediction variables in the discriminant function, then, provides no basis for confidence that they do represent important causal forces explaining differences in rates of economic growth or the potential for growth.

II. Regression Analysis

The regression stage of the model is an effort to explain the four prediction variables used in the final form of the discriminant function. These are introduced as the dependent variables in four multiple regression equations. Five additional regression equations are developed to explain some of the independent variables in the first set, and

four more are then developed to explain some of the independent variables in the second set.

In building an equation for each dependent variable, the authors apparently began with all twenty-eight other variables as candidates for inclusion and let the computer choose those most strongly associated with the dependent variable. In a few cases "variables were omitted from equations in which their inclusion would lead to a clear-cut violation of the direction-of-causality constraint," but the criterion was applied only "sparingly." The unfortunate result of this restraint is that very few of the actual regression equations appear to describe genuine causal relationships between independent and dependent variables. There are three general sources of doubt.

Definitional Overlap

Several of the dependent variables have definitions which seriously overlap those of the independent variables being used to explain them. For example, regression equation (6) finds "the character of agricultural organization" most strongly explained by the "size of the traditional sector." Yet "agricultural organization" runs from "communally owned agricultural lands in which the marketing of crops is only of incidental importance," up to large and viable "commercial owner-operated farms." The "traditional sector" is defined as "traditional subsistence agriculture in which . . . marketing of surpluses [is] of incidental importance"; it specifically excludes "modern commercial agriculture." Equation (7) in turn finds "the size of the traditional sector" to be negatively dependent on "the extent of dualism." Yet the low point of the "dualism" scale is "the largely agrarian society having an extremely small exchange sector." All we really know from these relationships is that there are many countries with a large share of the population in subsistence agriculture—whether this be classified as high on a "traditional sector" dimension or low on "dualism" or "character of agriculture" dimensions. If a high degree of correlation among

these dimensions holds any interest at all, it is only as a reliability check on the coding techniques.⁷

Implausible Causation

Many of the regression equations associate effects with very implausible "causes." One general difficulty is that variables representing short-run rates of growth are sometimes used to explain variables which express long-run levels. In equation (11) "the extent of literacy" (level for population aged 15 and over in 1958) is explained in part by "the rate of improvement in human resources" (measured by attendance ratio of school-age population in 1961). The two variables should be related—some countries have stronger traditions of mass education than others—but hardly causally, given that almost no one aged 15 or over in 1958 was of school age in 1961. In equation (12) the size in 1961 of the "indigenous middle class" is explained by the extent of social mobility in 1961.

A second difficulty is that some of the equations seek to explain very "hard" economic realities in terms of rather "soft" social or political rankings. For example, equation (8) explains the degree of "dualism" in the economy as partly dependent on "modernization of outlook." Equation (6) explains the "character of agricultural organization" partly in terms of "leadership commitment to development." Any implication that such transient attitudinal variables as "commitment" and "outlook" are significant causes of the size of the subsistence sector is better postponed to the day when the scientific bases of Couéism are more firmly established.

Inferences of Mutual Causation

Strong associations are far too readily interpreted as evidence of mutual causation. When two variables are strongly associated, it is not unexpected that each will figure prominently in a multiple regression equation for the other. In five cases where this occurs the authors do not hesitate to attrib-

⁷ For other examples, see Eckstein, pp. 17-18.

ute this result to "reciprocal causation" or, more grandly, to a "mutually reinforcing feed-back relationship." For example, the variable "rate of development of human resources" is found in the equation explaining the "change in the degree of industrialization," while the industrialization variable is found in the equation explaining the human resources variable. In the authors' interpretation, not only does human investment remove constraints on industrialization, but industrialization creates a necessary demand for human investment. The authors offer no justification for automatically discarding three alternative interpretations of the observed association: the unidirectional impact of industrialization on human investment, the unidirectional impact of human investment on industrialization, or the impact of a third variable (e.g., governmental activism) on both phenomena. Nor do they consider whether the regression coefficients should be adjusted downward from their observed levels so that each does not reflect individually a full association that they are assumed to be causing together.

III. *The Multipliers*

The culmination of the econometric model is the calculation of multipliers estimating "the relative impact of the various economic and noneconomic forces represented in the model upon the potential for economic development." Only the four prediction variables are assumed to have a direct impact on development potential; but all eighteen independent variables in the regression equations are assumed to have an indirect impact through the network of relationships to the prediction variables which those equations imply. There are two important problems with this multiplier analysis.

Utilization of Cross-Section Relationships

As econometricians have long recognized, there are substantial limitations on our ability to use cross-section analysis either to infer intertemporal relationships from the past or to project such relationships into the future. The inference of past time-series relationships requires that they have been

neither exaggerated nor obscured by such factors as: differences among countries in their economic institutions; differences in the environment—external and internal—within which different economies function; causal interactions among variables; disequilibria in relationships at the time of observation; and the impact of the relative performance of other nations. Even if the cross-section estimates actually do reflect intertemporal relationships, their use as a basis for projections into the future requires that no significant lags will exist in a country's response to a new level of performance; and that dynamic considerations—technological progress, changing consumer tastes, and a shifting structure of world demand at higher world income levels—can safely be ignored.⁸ The authors do not address themselves to any of these thorny issues. It seems unlikely, however, that these requirements are met by very many of the relationships described in their model.

Lack of Independent Basis

The authors never attempt to establish a logical relationship between the discriminant analysis and the multipliers. In the earlier analysis all twenty-nine variables were examined for their relationship to development potential, and four of these were found sufficient to account for more than 97 percent of the variance between groups. Are we to take that result as meaningful, or must we go beyond it and see as significantly related to development a whole slew of variables that were already tried and found unable to account for more than 3 percent of the remaining variance? If we must go beyond the results of the discriminant function, would we not be more interested in the *direct* association of each of the variables with development potential itself, rather than the indirect association through the four variables in the discriminant function?

These questions could have practical implications, because there are wide discrepancies between the results of two analyses.

⁸ See Eckstein, pp. 21–22, for possible applications to the model.

Two prediction variables—the degree of improvement of financial institutions and the degree of modernization of outlook—emerge from the multiplier analysis as still the most important influences on development potential. However, the other two prediction variables—leadership commitment to economic development and improvement in agricultural productivity—fall to eighth and tenth place, respectively. What advice, then, are we to give to the development planner who might wish to act on the basis of the values in the Adelman-Morris analysis? Should he seek to inculcate leadership commitment and to increase agricultural productivity, as suggested by the discriminant analysis? Or should he attempt to reduce dualism, to make physical overhead capital more “adequate,” to accelerate industrialization, and to increase investment in human resources—all of which have greater importance in the multiplier analysis?

Anyone tempted to employ the Adelman-Morris multipliers should examine the highly indirect relationships on which they are constructed. For example, the seventh largest multiplier is that for “rate of improvement in human resources.” Fully 61 percent of its value comes from making possible more rapid industrialization, and in turn fully 85 percent of the multiplier for industrialization comes from its impact on “improvements in financial institutions.” The latter variable carries the largest multiplier of all, and both the earlier ones are largely based on it. The implication then, is that one should attempt to influence a variable that has no directly-demonstrated relationship to development primarily because it may influence another variable that also has no directly-demonstrated relationship to development but that may influence a third variable which may have a direct relationship. It may be that “For want of a nail . . . a kingdom was lost,” but more than mere statistical association would be required to establish so tenuous a chain of causation.

IV. *Structured vs. Unstructured Models*

Throughout their joint work—including the factor analysis not covered in this dis-

cussion—Professors Adelman and Morris have relied on completely unstructured statistical techniques. In both the discriminant and the regression analysis, they “consciously avoided a priori specification of the functions” and instead presented the computer with a long list of variables and “let the data specify the model.” Thus, they are forced to judge variables entirely by their contribution to the explanatory power of the equation. The high multicollinearity in their data, however, renders this doubly dangerous. One typical result of collinearity is that coefficients of the variables have very large standard errors of estimate. Even when a coefficient is significantly different from zero (as are most of those in the Adelman-Morris regression equations), it is very unlikely to be significantly different from the coefficients of other closely-related variables. Thus, discriminant functions and regression equations are built on the bases of very narrow statistical differences, and their compositions are easily determined by random influences or small biases. Yet the authors take their results at face value and assume, for example, that the twenty-five variables not included in the discriminant function are important only insofar as they help explain the four variables that are included. When the data are highly collinear, however, the case is stronger than ever for using as much logic, theory, and prior knowledge as the investigators can muster in order to preclude a capricious selection among available explanatory variables.

The authors argue that their unstructured statistical techniques are necessary at this stage in the study of development, “since there are no firmly validated theories of the process of socio-economic and political change.”⁹ Yet even if we have very few answers—established quantitative measures of the empirical relationships among variables—certainly we know enough about the logical relationships among variables and

⁹ Note that the authors themselves must have had such a theory in mind when they required that a fast-growing country score high on five out of seven particular variables in order to qualify as having “high development potential.”

levels of analysis to be able to pose questions in an ordered manner. For example, the rate of growth of *GNP* per capita might be expressed as the rate of growth of *GNP* minus the rate of growth of population. The growth of total *GNP* might then be disaggregated into a weighted sum of either the growth of sectors or the growth of individual factors and their overall productivity. The growth of population might be expressed as the difference between crude birth rates and the rates of mortality and net emigration. Given these disaggregated measures of performance, development might be conceived as a three-stage process in which the country's "resources"—its natural wealth, physical capital, and human attributes—become translated through "economic action," governmental policies and private behavior, into "economic performance."¹⁰ The point, however, is not the operational or theoretical merit of a particular framework, but merely that the absence of "firmly validated theories" does not prevent us from conceptualizing a process in a way that may lend some order—and greater meaning—to empirical analysis.

V. Conclusions

There are many problems inherent in any attempt to separate out the strands of causation in the complex web of differences in the performance and potential of underdeveloped countries. Unfortunately, Professors Adelman and Morris have not adequately addressed themselves to these problems. There may be methods of direct observation that would permit the objective measurement of such subjective phenomena as "commitment" and "outlook"; but it is impossible to be confident that purely "judgmental" measures can ever be purged of the judge's preconceptions as to their causes and concomitants. Measures of the rate of growth of

per capita *GNP* in underdeveloped countries are themselves fraught with great uncertainty; but even the best of empirical measures would prove an inadequate foundation for the elusive concept of "development potential." Comparisons across countries can offer clues as to intertemporal relationships; but careful justification would be required before policy implications could safely be inferred from even the best-established of cross-section findings. No statistical technique can by itself separate out causes, effects, and joint effects or even the most important associations among closely correlated variables; but completely unstructured techniques, in which theory and prior knowledge are neither tested nor used to order the relationships are the least likely to produce meaningful results.

The greatest problems with the study, however, are not the inherent ones. In order to persuade us of the validity of their model as a whole, the authors would have to demonstrate first the soundness of both their discriminant and their regression analyses and then the legitimacy of their "multipliers" as a separate level of inquiry. Yet in neither statistical analysis do they establish either the independent measurement of the variables they are relating or the structural logic of the relationships they are describing. They also fail to make clear why the "multipliers," describing as they do indirect associations with development potential, should be taken more seriously than any direct associations that might be observed.

The general conclusion the authors offer is that "the important impediments to increasing capacity to develop" that emerge from the model "are social and political as well as economic." This is intrinsically a rather plausible statement. Unfortunately, neither its plausibility nor that of the specific relationships described has been augmented by any of the statistical results which comprise the "econometric model of development." On the contrary, our analysis of the model suggests that formidable obstacles remain to be confronted before the measurement of cross-sectional statistical associa-

¹⁰ This would imply that very unlike processes are being mixed in the authors' attempt to compare in one function the relative importance of two "resources" (modernization of outlook and leadership commitment), one set of "actions" (improvement in financial institutions), and one element of "performance" (improved agricultural productivity).

tions can begin to challenge economic analysis—the creative interaction of general theory and the knowledge of particular institutions—as the fundamental tool of policy-making for economic development.

REFERENCES

- I. Adelman and C. T. Morris, *Society, Politics and Economic Development: A Quantitative Approach*, Baltimore 1967.
- , "Performance Criteria for Evaluating Economic Development Potential: An Operational Approach," *Quart. J. Econ.*, May 1968, 82, 260–80.
- , "An Econometric Model of Socio-economic and Political Change in Underdeveloped Countries," *Amer. Econ. Rev.*, Dec. 1968, 58, 1184–1218.
- P. Eckstein, "Quantitative Measurement of Development Performance: A Critique of the Adelman-Morris Model," Discussion Paper No. 7, Center for Research on Economic Development, Ann Arbor, April 1969.
- M. G. Kendall, *A Course in Multivariate Analysis*, London 1957.

An Econometric Model of Development: Reply

By IRMA ADELMAN AND CYNTHIA TAFT MORRIS*

The comments of Peter Eckstein and Sarah Berry highlight the major methodological issues in our cross-section analyses of socio-economic and political change in underdeveloped countries. Our purpose in this reply is to clarify those characteristics of our research design and data which distinguish them from more conventional approaches; and to note our agreement or disagreement with the major comments. Our reply touches on both our earlier and current work since "An Econometric Model of Development" represents only one segment of our research and treats only sketchily the underlying methodological issues.

Before proceeding further, we would like to emphasize the limitations we have imposed on our work. We have not claimed to offer a "reliable guide to policy-making." We have not attempted to explain differences in rates of change of per capita *GNP*. We certainly have not claimed at this early stage of our inquiries to have "separated out causes, effects, and joint effects." We have used an empiricist approach to construct an exploratory model and have described cause and effect relationships *implied by the model*. We can make no better correction of Eckstein's view of our purposes than to repeat the final statement of our paper:

... it is to be hoped that the various noneconomic and economic features of low-income countries that appear *in the present model* to have particularly strong effects upon development potential may give some indication to social scientists of the most profitable directions for research into the determinants of intercountry variations in overall capacity to perform well economically in the long run.¹

* The authors are, respectively, professor of economics at Northwestern University and professor of economics at The American University.

¹ Adelman and Morris (December 1968, p. 1212). The italics have been added.

I. Statistical Methodology: General Comments

In designing our recent research on the interrelationship between economic and non-economic forces in economic development, we have used statistics as a tool for indicating the structure of the underlying phenomena rather than as a device for testing hypotheses. We have chosen this approach as most efficient in view of the extremely small quantity of validated knowledge concerning the "laws" governing the complex interactions between economic, social, and political forces in the process of economic development and modernization. Since the statistical theory for testing a given hypothesis against more than one alternative is at an early stage of development,² the hypothesis-testing approach is only efficient when an operationally testable hypothesis can be formulated, and the equally plausible alternative hypotheses capable of formulation are few in number and clearly distinguishable. Neither of these prerequisites appears to be satisfied in the field of research under discussion.

The literature of economics, history, sociology, and anthropology have produced a large number of reasonable, yet not well validated theories which would justify the choice of a multiplicity of alternative variables for the explanation of intercountry variations in economic development. Therefore we see no grounds for Eckstein's view that an a priori choice among competing hypotheses is either less arbitrary or more meaningful than a selection made by a reasonable and fairly robust empirical procedure. We need hardly add that we reject his recommendation that the problems posed by the

² For a discussion of this point, see E. L. Lehman who develops the theory of hypothesis testing against multiple alternatives for the case in which the alternative hypotheses can be parameterized. The latter condition is seldom met for the kind of alternative hypotheses relating to our present research.

presence of interrelated qualitative influences be solved by a priori specifications limited to directly measurable and (we suppose) weakly intercorrelated influences. It so vitiates the formulation of the problem that it in effect would require that we change the subject of our inquiry.

The only specific suggestion for an alternative model which Eckstein makes relates only marginally to the subject of our research. Furthermore, it by no means escapes the methodological problems posed by the existence of a multiplicity of interrelated influences in economic development. His three-stage model would face a serious problem of direction of causality since the private and government actions by which his resources are translated into economic development are, in fact, major causes of changes in the quantities of inputs. In addition, there are significant intercorrelations among the various actions as well as among the various elements of economic performance. Eckstein's failure to escape these difficulties reinforces our view that the problems he cites are inherent in our subject of inquiry and are not quite so amenable to solution as he supposes, even with a priori specified models.

It is our belief that there are many different ways to approach subjects of scientific inquiry, each with its own special defects and advantages as well as defects and advantages held in common with other approaches. We are not recommending the rejection of structured techniques nor objecting if others prefer to apply them to our subject. We simply consider them less suitable than unstructured techniques for our particular research given the general state of knowledge and our particular resources at this point in time.

The Robustness of our Research

Mr. Eckstein and Mrs. Berry have accused us of arbitrariness particularly with respect to decisions we have made in preparing our various definitional schemes. Yet they are both certainly aware that arbitrary assumptions do not, in and of themselves, invalidate a model. The first

point we would like to stress in discussing the arbitrariness of our assumptions is that the choices we faced were not, as our critics imply, between arbitrary decisions we have made and nonarbitrary decisions we might have made. On the contrary, because of the unsatisfactory state of the general theory of social change and modernization, the fact that the concepts in which most theoretical discussions are couched are not operationally defined, the inadequacies of factual information on developing countries, and the prohibitive cost of direct measurement, our choices of assumptions for data construction invariably involved a selection among arbitrary alternatives. In making the necessary decisions, we relied upon the same considerations which influence the decisions of most other investigators; that is, a priori reasoning, judgments regarding data availability, our common sense, and an evaluation of the sensitivity of the results to our choices. However, whereas the various simplifying assumptions on which conventional economic composites are based are familiar, there are few precedents for the choices which we had to make in order to prepare our definitional schemes.

In discussing the arbitrariness of our assumptions, Eckstein in particular has employed some fine rhetoric to indicate that a number of decisions made in the course of our research *could* affect the results we obtained. Yet nowhere does he even hint that the evaluation of the extent of arbitrariness of any assumption in any model requires, among other things, a study of the *actual* sensitivity of the results of the model to reasonable alternative a priori specifications. We would not, ourselves, claim to have avoided completely assumptions to which our results will prove sensitive upon further testing. Nevertheless, since neither Eckstein nor Mrs. Berry has performed any sensitivity tests, despite the fact that the raw material for doing so was available to them, their confidence that the results would vary greatly with specific alternative definitions, measurement of the variables, and choices of procedures is rather surprising.

For example, Eckstein contends that we

TABLE 1—COMPARISON OF FINAL MULTIPLIERS
RESULTING FROM THE USE OF DISCRIMINANT
FUNCTIONS D_1' AND D_2' ^a

Variable	Multipliers		Ranks based on Multipliers	
	D_1'	D_2'	D_1'	D_2'
X_{13}'	2.212	2.309	1	1
X_{31}'	1.705	2.169	3	2
X_{21}'	1.832	2.13	2	3
$X_8'^b$	1.545	1.98	4½	4
X_4'	1.545	1.624	4½	5
X_{24}'	.913	1.062	6	6
X_{22}'	.873	.955	7	7
X_{49}'	.727	.940	9	8
X_{20}'	.754	.884	8	9
$X_7'^c$.095	.823	13	10
X_{23}'	.685	.797	10	11
X_{10}'	.414	.536	11	12
X_{27}'	.347	.407	12	13
$X_{19}'^c$.033	.288	15	14
X_{17}'	.052	.018	14	15
X_{29}'	-.232	-.300	17	16
X_{14}'	-.253	-.310	18	17
X_6'	-.168	-.320	16	18

^a See text for explanation.

^b One needs to add to this the change in physical overhead capital

^c Agriculturally related.

ought to have used the first discriminant function (D_1') obtained in our original discriminant analysis rather than the second (D_2') as a starting point for our econometric model. Table 1 presents the multipliers obtained using both D_1' and D_2' . It indicates that the ranking and order of magnitude of the multipliers of the first nine variables are virtually unaffected. The primary effect is to reduce the already low multipliers of the agricultural variables while leaving the order of magnitude and ranking of the other multipliers essentially unchanged. This was to be expected, since the replacement of the index of improvements in physical overhead capital by the indicator of improvements in agricultural productivity is the principal way in which D_2' differs from D_1' . The finding that the final multipliers are fairly insensitive to alternative specifications of the structure of the present econometric model is further reinforced by the virtual identity of the multipliers obtained through the stepwise procedure with those which result from

the application of the techniques of canonical analysis to the same data. (See Adelman, Geier, and Morris, May 1969.) As is well known, canonical analysis is equivalent to a limited information simultaneous equation procedure. Thus, the ranking in the model of the important forces affecting development potential appears rather insensitive to the choices made in constructing both the discriminant function and the model.

Mr. Eckstein discusses certain variables in our discriminant function which he implies entered merely by chance. We have tested the robustness of the choice of variables in our final discriminant function to alternative statistical criteria for adding and deleting variables. An alternative stepwise discriminant procedure which uses the F -ratio test instead of the generalized distance criterion yields results identical to D_2' .

Mrs. Berry criticises the arbitrariness involved in our choice of scoring scale. In our reply to O. T. Brookins' criticism of our scoring scheme we presented factor analytic results based upon three alternative transformations of our numerical data inputs (logarithmic, squaring and rank), (see Adelman and Morris, forthcoming). Not only does the mean absolute difference in factor loadings never exceed .08, but in only one instance out of 267 for the three test analyses is the loading sufficiently different to lead to the assignment of a variable to a factor different from that to which it originally was assigned.

How many conventional econometric models perform as well? In different estimates of consumption functions with similar data but over different (though overlapping) time periods, the marginal propensity to consume out of quarterly disposable income was found to vary between .1 and .6, the long-run marginal propensity to consume to range from .60 to .98, and the estimated coefficient of past peak income to be anything from a significant value of -.625 to an insignificant value of -.241. (See Griliches, Maddala, et al.; Zellner; Duesenberry et al.) In alternative estimates of production functions for the United States the elasticity of output with respect to labor was estimated

to be as different as .18 and 0.95 despite the use of the same functional form. (See Klein and Goldberger, Valavanis-Vail.) In the study of investment behavior, estimates of the elasticity of capital stock with respect to relative price range anywhere from a value of .8 to a value of .09, and with respect to output anywhere from a value of .6 to a value of 1.0. (See Eisner.) Yet, all these widely divergent estimates were obtained by well recognized econometricians sharing a common theoretical framework and applying similar econometric techniques to analogous bodies of empirical information. Thus, a comparison of the sensitivity of the results of our analyses with the sensitivity of results of other econometric work indicates that our research is quite robust, not only in absolute terms but also by the standards of more conventional econometric studies.

Since we have used unstructured techniques of statistical analysis for the purposes of hypothesis-making, there is yet another test of validity applicable to our research, namely, whether the hypotheses suggested by the various statistical manipulations of our data are mutually consistent. After all, the use of inductive techniques is likely to provide a good foundation for theory construction only if the hypotheses suggested by different analyses are both sensible when viewed in isolation and mutually compatible.

While it is somewhat too early in our work to apply this criterion of robustness with great stringency, we can cite several general propositions whose validity is reinforced by the several different pieces of empirical analyses. First, our work shows that in the analysis of economic development, the division between endogenous and exogenous variables along traditional disciplinary lines can produce misleading results. This conclusion emerges from our socio-political factor analyses published in 1967 in which social and political forces explain 79 percent of intercountry differences in economic growth rates for low-level underdeveloped countries, 49 percent for the intermediate-level, and 61 percent for the high-level (pp. 187, 213, 241). It is reinforced by our discriminant analysis of May 1968 in which only 2 out of 4 vari-

ables explaining variations in development potential are economic (p. 277). It is further strengthened by our paper in this *Review* in which we found that 67 percent of the relationships of the economic variables are with noneconomic variables and that, of the multipliers which exceed .80, only 4 are purely economic (p. 1201 ff).

Second, in studying economic performance, one should distinguish among a minimum of three different typologies of underdevelopment. In our factor analysis of 1967 we found systematic differences between the patterns of interactions among economic and noneconomic influences upon rates of economic growth for the three subsamples representing successive levels of development (pp. 6-7, and Chaps. V-VII). This evidence is reinforced by our canonical analysis of May 1969 in which we found systematic variations in relationships between policy instruments and policy goals for countries at different levels of development (pp. 416-23).

Third, for an average underdeveloped country, the connections between economic and sociological forces are more intimate than the connections between economic and political forces. In our 1967 full-sample factor analysis of per capita *GNP*, 40 percent of intercountry differences in income per head were explained by social and socio-economic variables and only 26 percent by political forces (p. 151). In our econometric model in this *Review*, 90 percent of the interconnections between economic and noneconomic variables were between variables in the economic block and variables in the social and socio-economic blocks and only 10 percent were with political forces (pp. 1201-3).

Fourth, the most important political variable for economic performance is the nature of leadership attitudes towards economic development. This conclusion emerges from our 1967 socio-political factor analysis of rates of economic growth for high-level underdeveloped countries in which extent of leadership commitment by itself accounted for approximately 59 percent of intercountry differences in rates of economic growth. It is reinforced in our May 1968 discriminant analysis in which leadership commitment in

economic development was the only political variable in the discriminant function. This conclusion is also emphasized by the multiplier results of our econometric model in which the extent of leadership commitment to economic development was the only political variable to enter the "top ten" list.

Thus there would appear to be a fair extent of congruence among hypotheses derived from various different statistical analyses performed by us with our data.

We have been quite aware that the appropriate scientific procedures for evaluating a model include fairly extensive testing of the sensitivity of the results to reasonable changes in its specifications. Consequently, an important part of our past and present research has been devoted to such tests.³ Preliminary tests suggest that our results are not very sensitive to reasonable variations in the specification of our

³ In addition to the sensitivity tests reported upon in the text, while preparing our data we tested the sensitivity of our factor analyses to changes in the definition of selected variables for which we had a choice of feasible possibilities. In defining the degree of centralization of political power in two different manners, for example, we found our results substantially unchanged. Compare the factor analysis results based on the first definition in our 1965 paper (p. 562) with those based on the second definition in our 1967 study (p. 58 and p. 151). We defined improvements in human resources in several ways: one stressing primary education; one emphasizing primary education for low-level countries and secondary education for more developed countries; and one based on a combined index of secondary and higher education. Since our results for the different samples did not change significantly with these variants, we chose the third alternative as the simplest to prepare and as reasonable a priori as the others.

We have also tested the sensitivity of our factor analyses to changes in the loss of information due to our grouping procedures by varying the number of categories for several indicators (e.g. changes in per capita GNP, gross investment rates, crude fertility rates) for which we had point estimates for the majority of countries in the sample. Since the results proved relatively invariant to changes in the number of class intervals, we chose that number of brackets which gave us both reasonable discrimination among countries and reasonable reliability of the judgmental information used to make interval estimates for countries without point data.

data,⁴ to our scoring techniques, or to the detailed structure of our econometric model. In addition, different statistical approaches appear to lead to mutually reinforcing conclusions.

II. *Data: Theory and Taxonomy*

Our research efforts of the past six years have been directed to the interaction of non-economic and economic influences in the economic development of underdeveloped countries. We have chosen to approach this subject through quantitative analyses of those social, political, and economic characteristics of underdeveloped countries which, according to a priori reasoning, descriptive historical studies, and the relevant social science literature seemed likely to affect the economic performance of developing countries. Our subject of inquiry and choice of approach required us, therefore, to devise measures of qualitative non-economic and economic institutional characteristics for which the full range of underdeveloped countries could be ranked.

Direct measurement of the relevant influences was rarely possible because of overwhelming data deficiencies. Choices of indirect measurements were constrained by the small number of feasible alternatives for the measurement of each influence. Finally, in choosing to include countries for which all data were poor (usually those at the lowest development level) and thereby to avoid a major bias present in most cross-section studies of developing countries, we were restricted to measurement by intervals sufficiently wide so that descriptive and judgmental information could be used to make reasonably reliable interval estimates.

Our critics have failed to distinguish clearly

⁴ As would be expected, our results have proved sensitive to changes in definition which involved major conceptual revisions. Our original measure of "intensity of nationalism and sense of national unity" associated with leadership characteristics. We replaced it in response to criticisms by Everett Hagen with a measure of the "degree of national integration and sense of national unity" which associated with indicators of social development. Compare the results cited in fn. 3. (See also p. 55 of our 1967 study.)

between difficulties peculiar to our data and those which our data share with conventional economic and psychological data. We shall therefore discuss what are the essential attributes of our data. In particular, we shall point out their multidimensionality, their qualitative and ordinal character, and our translation of ordinal ranks into cardinal scores. We will then consider in some detail the conceptualization and definition of qualitative indicators.

Multidimensionality of Data

Almost all our variables are composites of several component elements. Consequently, as for all composites including such indices as *GNP*, cost of living, and intelligence, either a formal theory or an a priori conceptualization of the phenomenon is required in order to construct an aggregate index. What distinguishes our data from conventional aggregates is not the use of theory in measurement, but the fact that the a priori judgments we have made in aggregating the components of each index are less securely based in explicit theory than is the case for conventional composites.

Qualitative Character of Data

Most of our indicators are either purely qualitative or partly qualitative variables in which descriptive information and expert opinion were used to classify individual countries. The ranking of individuals by expert opinion is a completely accepted procedure in psychology; nevertheless, there are several pitfalls in its application. Experts may interpret concepts and definitions differently; their opinions about the facts may differ; and they may be biased. We attempted to minimize the effects on our data of these sources of variation in expert opinion; we used interviews to probe differences in expert reactions to both our definitions and the individual country classifications. Moreover, we attempted to define our concepts operationally prior to the consultation of experts. By using relatively objective criteria for ranking individual countries, we sought to eliminate systematic errors in our

data arising from expert biases of various sorts.⁵ With respect to the extent of leadership commitment for economic development, for example, these criteria included: the presence or absence of cooperation between the government ministries, planning agencies, and national banks, engaged in central guidance of the economy in actions to promote growth; the presence or absence of leadership measures for institutional changes to promote economic growth (e.g., land reform); the presence or absence of full-time planning groups engaged in both planning for and execution of national plans. The application of these criteria permitted us to rank the majority of countries with the use of descriptive information. Experts were consulted only in order to resolve doubtful cases. Our critics would do well to study carefully the criteria for classifying countries specified in our actual definitional schemes. Mr. Eckstein appears to have based his criticism on his intuitive view of the concepts suggested by the titles of the indicators, while Mrs. Berry has based her criticisms on her reactions to our general statements preceding the definitions.

While it is obvious that we have not succeeded in eliminating all expert biases of various sorts, we reject Eckstein's contention that there is a *systematic* bias in our data due to the tendency of experts to assume that good growth performance is achieved by committed leadership. The evidence supporting our view⁶ is that the simple correlations between rates of growth of per capita *GNP* and most of our socioeconomic and political variables are significantly different for the subsamples representing different levels of development; the correlations between economic growth rates and leadership commitment to development, for example, are .59, .16, and -.02, for the high, intermediate, and low groups of countries, respectively. To maintain that the

⁵ For a useful discussion of the nature and varying importance of different kinds of errors in measurement, see our 1967 study (pp. 199ff.)

⁶ It is readily available in the Appendix of our 1967 study.

barely statistically significant correlation of .35 for leadership for the full sample is due to bias, it is therefore necessary to assert that the bias of experts varies significantly with the level of development—a rather implausible contention. To reemphasize our basic point, we do not claim that some or even many experts were not biased in various ways but only that these biases do not appear to have systematically distorted our results.

Cardinal Translation of Scores

The third special characteristic of our data, the assignment of numerical scores to ordinal data is criticized by Mrs. Berry. We have discussed this in detail in our forthcoming paper, but let us note here again that our procedure is an accepted and conventional one in the field of psychology. It is no more arbitrary than a host of other commonly accepted econometric simplifications and abstractions. The practical point, surely, is the sensitivity of the results to the choice of scale; as already indicated, our preliminary tests suggest that the structure of interrelationships underlying our data are surprisingly insensitive to wide variations in methods of scoring.

It should be noted that the choice of a cardinal scale with equal intervals even for a measurable characteristic such as school enrollment ratios is in no way free of possible bias. It involves a selection among such alternatives as raw scores, squares, logarithms or other transformation. This choice must be based in part on the presumed relationships between school enrollment ratios and whatever underlying characteristic (such as the quantity of skilled human resources) they are presumed to measure as well as on the presumed relationship between that underlying characteristic and the purpose of the investigator's inquiry. In our own data preparation, a priori reasoning together with the other sources of knowledge available to us were used in order to choose breaking points between intervals.

The Conceptualization of Qualitative Indicators

The major difficulties of our data lie, in

our opinion, neither in the fact that they are qualitative and based on judgmental information nor in our translation of ordinal ranks into cardinal scores. The core difficulties lie rather in the conceptualization and definition of phenomena for which there are no theories as precise and acceptable as those on which qualitative economic measures are based. There are few available theories suited to defining continua along which the full range of underdeveloped countries can be ranked for most of the characteristics in our study. In addition, there are little data to fit such theories as are current. Furthermore, our efforts at definition were constrained by the frequent necessity to use indirect rather than direct evidence. As generally recognized, the use of indirect evidence requires inferences concerning "presumed connections, usually causal, between what is directly observed and what [a] term signifies . . ." (Abraham Kaplan, p. 55).

The process of conceptualization and definition which we followed in constructing our composite indicators is a procedure well tried in the history of scientific inquiry—in the physical as well as the social sciences. Abraham Kaplan in *The Conduct of Inquiry* has the following to say about the derivation of scientific concepts in the early stage of scientific inquiry:

In short, the process of specifying meaning is a part of the process of inquiry itself. In every context of inquiry we begin with terms that are undefined—not indefinables, but terms for which that context does not provide a specification. As we proceed, empirical findings are taken up into our conceptual structure by way of new specifications of meaning, and former indications and references in turn become matters of empirical fact . . .

What I have tried to sketch here is how such a process of "successive definition" can be understood so as to take account of the openness of meaning of scientific terms. *For the closure that strict definition consists in is not a precondition of scientific inquiry but its culmination.* To start with we do not know just what we mean by our terms, much as we do

not know just what to think about our subject-matter. We can, indeed, begin with precise meanings, as we choose; but so long as we are in ignorance, we cannot choose wisely. It is this ignorance that makes the closure premature. [p. 77-78, *Italics added*]

In the preparation of our qualitative multidimensional indicators, we began with a priori definitions. Next, we studied the descriptive data in order to see how well actual country situations fit our formulation of the concept. The inadequacies of the initial fit were then used to reformulate the concept to fit better the characteristics of the real world. We then consulted expert opinion and again reformulated the definitions. We continued this process of confronting successive reformulations with information on actual country situations until we were able to classify the 74 countries in our sample with reasonable confidence.

Example of the Dualism Indicator

The construction of our indicator of the extent of socioeconomic dualism illustrates well the interaction of conceptualization and testing against the actual world which took place in the preparation of our data. We choose this indicator for discussion because Mrs. Berry criticizes it in her comment.

In constructing our dualism indicator, our major conceptual difficulty was the ranking of two intermediate categories, for we characterized both ends of the continuum by the absence of marked dualism. Dualism is absent at the lower end by reason of the overwhelming predominance of nonmarket subsistence agriculture combined with the extremely limited growth of a market-oriented sector; while the upper end is not markedly dualistic because the intermingling of modern and traditional elements throughout the economy resulted in the absence of a clearcut geographic cleavage between a market-oriented sector and a distinct predominantly nonmonetized traditional sector. In our initial a priori definition, we conceived of a single intermediate category characterized by sharp social, economic, and technological contrasts between a geo-

graphically distinct and important plantation, extractive, or industrial sector using advanced technology and a major subsistence nonmarket agricultural sector.

When we confronted our original a priori definition with descriptive data from country studies, we immediately discovered that it made no provision for countries with a geographically quite distinct, important, and rapidly growing market sector characterized by the predominance of indigenous cashcropping with conventional techniques. The countries in which peasant cashcropping dominate the market sector differ significantly from those in which extractive or plantation enterprises using advanced technologies dominate the market sector. In particular, resource flows in the latter countries are limited primarily to the intermittent labor flows, while in the former countries, resource flows included movements of land, labor, and capital.

In reformulating the dualism indicator we therefore gave weight to the extent of intersectoral resource flows as well as to socioeconomic and technological contrasts and to the presence or absence of a relatively clearcut cleavage between market and nonmarket sectors. We defined two intermediate categories for countries with both a geographically distinct and important market sector and a relatively large predominantly nonmonetized traditional sector. We gave higher scores to countries in which: (a) the market sector was characterized by the predominance of indigenous cashcropping; (b) less marked contrast existed between technologies and styles of life; (c) more extensive economic interaction occurred between traditional and modern sectors. We did so because of our judgment that these countries were further along the path to the pervasive intermingling of modern and traditional elements throughout the economy characteristic of the upper end of the spectrum summarizing the extent of dualism.

The reformulated four-way classification scheme enabled us to rank the great majority of countries in our sample. An examination of additional sources of country information led to only minor reformulations of

the category descriptions which were necessary to take better account of the diversity of individual country situations.

In criticizing our dualism indicator, Mrs. Berry did not mention the actual definitional scheme which consists of the set of descriptions of the categories just discussed. Consequently, her attack on individual country classifications is most misleading. She does not state anywhere that her criticism is primarily in terms of a criterion, the extent of foreign finance and direction, *which is not even mentioned in the definition*. It is true that in our general comments we point to foreign finance and direction as typical of countries in the lower-ranking categories of dualism. However, because of those country situations Mrs. Berry cites, we explicitly did not include foreign finance and direction as a classificatory principle in the definitional scheme.⁷

On the general subject of the reasonableness of our individual country classifications, it is self-evident that, in making 3,034 classifications, we have probably made some misclassifications which will be pointed out to us in the course of time. It is nevertheless our opinion that the number and size of these errors are significantly less than in conventional quantitative economic data on underdeveloped countries.

⁷ The particular countries Mrs. Berry mentions as all having foreign financed and directed sectors can be classified reasonably well in terms of the criteria given in the definition. To illustrate with our ranking of South Africa, Rhodesia and Uganda: according to our information, South Africa was the only one of the three which in 1960 did not have at least a moderately definite cleavage between an important industrial and/or agricultural exchange sector and a single relatively large geographically distinct predominantly nonmonetized sector; it is consequently the only one of the three which met the stated criteria for category a. The minus score to South Africa was given because of the existence of relatively small predominantly nonmonetized agricultural pockets in its economy. Rhodesia and Uganda differed in 1960, according to our information, because the former had a significantly more important industrial, mining, and agricultural exchange sector than the latter and a relatively less important geographically distinct predominantly nonmonetized traditional sector. On both counts, Rhodesia was placed in category b rather than c.

III. *The Discriminant Function*

Mr. Eckstein's criticisms of our discriminant analysis consist of an attack on the original grouping by development potential; objections to our procedures in obtaining the final discriminant function; criticisms of the variables in the function; and criticisms of our interpretations. We will consider each in turn.

Grouping by Development Potential

The focus we chose for our discriminant and regression analyses was economic development in the broad sense of institutional transformations which create a capacity for widespread and continuous economic growth. For our purpose, the use of per capita *GNP*, however precise, was inappropriate since, as is well known, the raising of the growth rate of per capita *GNP* can take place, even over periods of a decade or so, without being accompanied by fundamental institutional change in other than a sporadic and limited way. In choosing to devise a less precise but more appropriate classification of countries, we considered that, *for our purpose*, a "gain" of information on the breadth of recent economic change would more than compensate for any "loss" of information on per capita *GNP*.⁸

The classification we devised is akin to a more recent effort by Paul Clark and Alan Strout to construct measures of economic performance of developing countries. To quote their recent paper:

The concept of a developing country's economic or development performance can cover many different phenomena. The growth process consists of a complex and convoluted series of cause and effect relationships ranging from plans and rhetoric to eventual improvements in the level and conditions of living. Performance can apply to the effectiveness of pursuing either intermediate or ultimate goals and can be observed at a wide variety of points in the development process.

⁸ It should be noted that we had already devoted an entire book to a factor analytic study of intercountry differences in rates of growth of per capita *GNP*.

Thus performance can include such diverse aspects as (a) choice of effective policies to promote desired growth and development; (b) effective implementation of policies chosen; (c) mobilization of additional resources for growth and development; (d) efficiency of use of both domestic and foreign resources; (e) structural and other changes necessary for longer-run economic, political, and social growth; (f) final effects on level and distribution of income and welfare. [p. 3]

In grouping countries by development potential, we sought, as do Clark and Strout, to distinguish between countries in which registered increases in per capita *GNP* were narrowly based and those in which they were accompanied by widespread economic change. Our three-way classification of countries by development potential would rank countries somewhere between the two ranks (by growth and policy performance) on the Clark and Strout index since our grouping is based on both kinds of performance criteria.⁹

We agree with Eckstein that independent evaluations of country development potential comparable to those used, for example, by Clark and Strout would be preferable to the wide range of indirect evidence on which our classification by development potential was based. However, because of the breadth of information we took into account, we see no particular reason to expect that intensive field research would produce country rankings very different from those we obtained.

Procedures for Obtaining Discriminant Function

The iterative procedure we used to obtain our second discriminant function is attacked by Eckstein as unsound. We disagree. Our aim, as explained in our original article, was to seek the best means for classifying countries into performance groups using a small number of performance characteristics. The procedure we followed is that developed by E. Forgy for deriving the optimal partition among a group of objects. This procedure

involves starting with a partition (in our study, we used the best partition we could devise on the basis of the external evidence on country performance available to us) and then reassigning each object to the group to whose center of gravity it is the closest. In Forgy's procedure, the value of the newly formed partition with higher value no longer results.¹⁰

The Variables in the Discriminant Function

We can hardly disagree with several of Eckstein's comments on the limitations of the variables contained in the discriminant function. First, we certainly have *not* obtained "independent causes of faster growth rates." The role of the variables in the discriminant function was merely "to classify countries into performance groups, using a relatively small number of performance characteristics." Second, in both the present article and the original discriminant study, we ourselves have been at pains to emphasize the fact that statistical associations may represent causality in either direction or be the result of common forces affecting both dependent and independent variables. We would, however, be fascinated to find any meaningful statistical analysis in which common underlying forces do not influence both sides of the equation.

Third, as we pointed out in both the present article and in our May 1968 article on the discriminant analysis, each variable in the discriminant function certainly represents closely related influences not directly measured by the variable itself as well as those directly represented. This possibility is surely self-evident for any statistical analysis in which the influences measured are proximate ones only. In our view, the appropriate method for understanding better the influences represented by a particular variable in a set of statistical results is to analyze carefully those variables which are the next-best alternatives to the particular variable chosen at each step in the analysis.

⁹ Exact comparisons cannot be made because the time period for their measures is different from that for ours.

¹⁰ For a good general discussion of the problems and procedures involved in deriving optimal classifications of *N* objects into categories, see H. P. Friedman and J. Rubin and the references cited therein.

We have recently rerun the discriminant function with a stepwise program having an F -ratio criterion for the entry and rejection of variables and giving F -ratios for all variables at each step.¹¹ The index of modernization of outlook is consistently the first variable to enter the discriminant function. As would be expected, the variables which are the next-best alternatives (as indicated by high original F -ratios combined with relatively large decreases in F -ratios as a result of the entry of the modernization index) represent a complex of closely interrelated economic and social influences, including the level of development of financial, agricultural and industrial structures, the extent of social mobility, the importance of the indigenous middle class, and the extent of secondary and higher education. Thus, the index of the modernization of outlook is undoubtedly closely related to basic economic influences.¹² Nevertheless, since the program was free to select direct economic measures and did not do so, it seems reasonable to interpret the choice of the modernization index as indicative of some independent influence of the noneconomic forces represented by the measure itself. With as many as 74 observations, and a difference of 1.3 in F -ratio, there is no reason to suppose that the choice of the modernization index rather than the leading economic index was the outcome of purely random influences.

The fourth point about the variables in the discriminant function on which we agree with Eckstein is that a positive relationship between improvements in agricultural productivity and growth performance (actual or potential) was to be expected since agriculture forms a very important sector in less-developed countries.¹³ One would simi-

larly expect improvements in industrial technology to be closely related to growth performance in advanced economies. But, what is the point? Is one to eliminate variables for which there is a strong a priori reason to expect a high intercorrelation? Is not an identification of such variables the main function of a priori theorizing?

IV. *The Stepwise Regression Model*

We now turn to the specific criticisms by Mr. Eckstein and Mrs. Berry of the nature of the stepwise procedure we employed and of the interpretation of our stepwise results.

Statistical Procedure

The particular stepwise regression procedure we applied is the one recommended by Draper and Smith for deriving *optimal* empirical relations from a body of data. Specifically, at each step in the procedure, a reexamination is made of the contribution of each variable in the equation as if it had been the most recent variable entered. This is done by computing the F -ratio for each variable in the regression and comparing it with a preselected percentage point of the appropriate F distribution. Any variable whose contribution is no longer significant is removed from the equation. Thus, the regressions obtained represent the best, in a least squares sense, summary of the multivariate associations present in the data. Despite Fisher's dictum against "fishing expeditions," this use of regression analysis to indicate and summarize the underlying multivariate associations seems to us a perfectly legitimate and technically more powerful extension of such measures of two-by-two associations as contingency tables and correlation matrices which are often used to investigate the presence or absence of interactions.

Interpretation of the Model

The major problems of interpretation raised by our critics are (1) the problem of direction of causality, (2) the problem created by the fact that any variable entering a

¹¹ Unfortunately, no program giving information on all variables is available for the stepwise discriminant procedure which applies the criterion of the greatest contribution to explaining the generalized distance between group means.

¹² See Adelman and Morris (1967, row 4 of Appendix Table A-1) (correlation matrix for social, political and economic indicators: full sample).

¹³ The simple correlations between rates of change of per capita GNP and improvements in agricultural productivity for the high, intermediate, and low samples,

respectively, are .78, .28, and .44. Thus the correlation is higher for countries with smaller agricultural sectors.

statistical equation may be a surrogate for influences other than those directly measured by the variable, and (3) the problem of interpreting cross-section results as representing historical processes.

Our original decision to apply sparingly the direction-of-causality criterion for omitting variables was based, as we stated in the paper, on the general state of ignorance regarding the potential interactions among the variables included in our study. We nevertheless agree that greater restrictiveness is desirable and have rerun the stepwise analyses with more restricted lists of potential causal indicators established by a priori reasoning based on an extensive examination of the relevant social science literatures. These lists do not, however, exclude cases of genuine mutual interaction since the exclusion of an interacting independent variable produces results which are biased by the omission of relevant influences.

We reject Eckstein's implication that an a priori selection among closely related causal variables yields a meaningful solution to the problem posed by variables which are proxies for influences other than those directly represented. The only effective method for understanding these influences is to study explicitly at each step in the analysis the *F*-ratios of the omitted variables, the net correlations between the dependent and the omitted variables, and the pattern of actual country residuals. These pieces of information and the matrix of simple correlations together can give a reasonable idea of the forces for which each variable in the regression equations may be presumed to stand.

We agree with Eckstein and stressed in our book (pp. 265-66) that, strictly speaking, cross-section analyses cannot be interpreted to have any specific time dimension. Consequently, an important part of our current research consists of cross-section studies for several different historical periods and individual country time-series studies of the process of economic development.

V. Conclusion

In conclusion, we firmly believe that the

empirical approach to scientific inquiry forms an extremely useful *complement* to the usual structured techniques of the economist for the study of the process of economic development. We reject the view that, in the investigation of the complex interactions among noneconomic and economic influences in development, a priori models are likely to produce much more meaningful results than models based upon empirically determined regularities. The multiplicity of plausible hypotheses regarding causes of variations in the capacity of low-income countries to perform economically is so great and the body of validated knowledge so small that a priori choices among candidate variables are inevitably quite arbitrary and tend to reflect primarily the disciplinary preferences of the investigator. We do not deny, and indeed are ourselves convinced that models with a priori specified functions are very important for the analysis and understanding of particular subsets of interactions about which a reasonable amount is known.¹⁴ We only propose strongly that alternative empirical procedures may be more fruitful for the initial exploration of those wider interactions involved in economic development which, by crossing disciplinary lines, involve relationships about which very much less is known.

The evaluation of any methodology involves an estimate of its success in achieving its purpose. As we stated clearly in the introduction of the paper under discussion, the purpose of our work has been to provide a good starting point for a process of experimentation in which further empirical testing and theoretical reasoning must interact. This we believe that we have done in view of the demonstrated robustness of our research and in view of the mutual compatibility of the hypotheses which emerge from our different analysis.

REFERENCES

- I. Adelman, M. Geier, and C. T. Morris, "Instruments and Goals in Economic Develop-

¹⁴ These subsets are usually encompassed by a single discipline.

- ment," *Amer. Econ. Rev. Proc.*, May 1969, 59, 409-26.
- I. Adelman and C. T. Morris, "An Econometric Model of Socio-Economic and Political Change in Underdeveloped Countries," *Amer. Econ. Rev.*, Dec. 1968, 58, 1184-1218.
- , "A Factor Analysis of the Interrelationship Between Social and Political Variables and Per Capita Gross National Product," *Quart. J. Econ.*, Nov. 1965, 69, 555-78.
- , "Performance Criteria for Evaluating Economic Development: An Operational Approach," *Quart. J. Econ.*, May 1968, 82, 260-80.
- , *Society, Politics, and Economic Development*, Baltimore 1967.
- , "Performance Criteria for Evaluating Economic Development Potential: A Reply," *Quart. J. Econ.*, forthcoming.
- P. Clark and A. Strout, *Aid, Performance, Self-Help, and Need*, Agency for International Development Discussion Paper No. 20, Washington 1969.
- N. R. Draper and H. Smith, *Applied Regression Analysis*, New York 1966.
- J. S. Duesenberry, O. Eckstein, and G. Fromm, "A Simulation of the United States Economy in Recession," *Econometrica*, Oct. 1960, 28, 749-809.
- R. Eisner, "Investment and the Frustrations of Econometricians," *Amer. Econ. Rev. Proc.*, May 1969, 59, 50-79.
- E. Forgy, "Cluster Analysis of Multivariate Data: Efficiency vs. Interpretability of Classifications," *Biometrics*, No. 3, 21 (abstract)
- H. P. Friedman and J. Rubin, "On Some Invariant Criteria for Grouping Data," *J. Amer. Statist. Assn.*, Dec. 1967, 62, 1159-78.
- Z. Griliches, G. S. Maddala, R. Lucas, and N. Wallace, "Notes on Estimated Aggregate Quarterly Consumption Functions," *Econometrica*, July 1962, 30, 491-500.
- A. Kaplan, *The Conduct of Inquiry: Methodology for Behavioral Science*, San Francisco 1964.
- L. R. Klein and A. S. Goldberger, *An Econometric Model of the United States 1929-1952*, Amsterdam 1955.
- E. L. Lehman, *Theory of Testing Hypotheses: Lectures*, Mimeographed Notes by C. Blym, Berkeley, no date.
- S. Valavanis-Vail, "An Econometric Model of Growth, U.S.A. 1869-1953," *Amer. Econ. Rev. Proc.*, May 1955, 45, 208-21
- A. Zellner, "The Short-Run Consumption Function," *Econometrica*, Oct. 1957, 25, 552-67.

The Effect of Income on Delinquency: Comment

By JOHN C. WEICHER*

In the March 1966 issue of this *Review*, Belton Fleisher proposed an analytical framework for assessing the relationship between income and juvenile delinquency and has attempted to measure the effect of income on delinquency. In a series of multiple regression analyses, he finds a "... general pattern of behavior ..." in which "the overall effect of income on delinquency appears to be negative ..." although there is a partial off-setting effect as a result of "... the positive influence of income on the payoff for property crimes." He also finds that "unemployment appears to be a cause of delinquent behavior ..." (p. 132), which is consistent with his previous work on the subject.

Fleisher's results, however, depend heavily on the choice of variables used to represent "tastes" for delinquency. It is the purpose of this comment to show that, at least for the only set of published data used by Fleisher, which is also the data regarded by him as most useful for investigating delinquency (74 community areas within Chicago): a more accurate measure of one important taste concept is available, and its inclusion substantially alters Fleisher's results. Secondly, Fleisher's economic variables are highly collinear with additional variables, reflecting other aspects of the taste for delinquency, which might reasonably be included in his regressions. When these variables are added, Fleisher's results are further modified and his conclusions, particularly those with policy implications, are vitiated.

I. Income, Tastes, and Delinquency

Fleisher employs three taste variables in his analysis of community areas within Chicago: *MOBILR*, the proportion of the population over five years of age that lived somewhere other than the community of current (1960) residence in 1955; *SPDVFM*,

the proportion of females over 14 years of age who are separated or divorced; and *NWHITE*, the proportion of the population that is nonwhite.

SPDVFM occupies a central role in Fleisher's analysis. It "... is supposed to represent the proportion of broken families in the community" [p. 124], a phenomenon that is important in part because children in broken families receive less parental supervision and guidance, and hence are likely to be more delinquency-prone. However, the proportion of children not living in broken families in the community can be calculated directly from Census data, as the ratio of persons less than 18 years of age who are living with both parents, to all persons under 18 (this variable will be subsequently abbreviated as *FAMILY*). The necessary data for this variable have been summarized for Chicago community areas in Kitagawa and Taeuber (see Table IV-1). *FAMILY* is more precise than *SPDVFM* in several respects. The latter includes childless separated and divorced women who have no potential delinquents to supervise; and it excludes widows, for whom the problems of supervision may be equally important. Further, as pointed out by Fleisher, each parent plays a part in supervising children; absence of either parent will reduce the extent of child supervision. On these grounds, *FAMILY* is a priori superior to *SPDVFM*.

The statistical effect of substituting *FAMILY* for *SPDVFM* is shown in Table 1. Equation (1-1) reproduces Fleisher's "All 74 Communities" regression (Table 5, p. 130). Equation (1-2) contains *FAMILY* instead of *SPDVFM*. *FAMILY* is statistically significant, in the sense that the term is used by Fleisher,¹ with the expected sign, while

¹ Fleisher employed a one-tailed test of significance at the 20 percent level to measure statistical significance (p. 127). In this comment, the usual two-tailed, 5 percent level will be used to measure statistical significance, unless otherwise indicated.

* The author is assistant professor of economics at The Ohio State University.

TABLE 1—ALTERNATIVE REGRESSION MODELS USING JUVENILE DELINQUENCY AS DEPENDENT VARIABLE—74 CHICAGO COMMUNITIES

	MEINC2 ^a	MEINC4 ^b	UNEMMC ^c	MOBILR ^d	NWHITE ^e	FORSTK ^f	SPDVFM ^g	FAMILY ^h	DILAP ⁱ	DENSE ^j	MSCOOL ^k	r ²
(1-1)	-.0051 (-2.8) ^a	+.00038 (+1.9)	+187 (+2.3)	+160 (+3.5)	+2.69 (+0.4)	+22.5 (+0.3)						.850
(1-2)	-.0059 (-2.3)	+.00016 (+0.5)	-46.6 (-0.5)	+103 (+2.4)	+8.96 (+0.8)			-5.79 (-1.3)				.868
(1-3)	-.0105 (-4.9)	+.00053 (+1.9)	-42.6 (-0.5)	+84.0 (+1.9)	+7.34 (+0.6)	-39.5 (-2.6)	-64.2 (-0.9)					.880
(1-4)	-.0073 (-2.9)	+.00013 (+0.4)	-114 (-1.3)	+65.2 (+1.5)	-11.4 (-0.9)	-42.6 (-2.9)		-6.61 (-1.6)				.883
(1-5)	-.0081 (-3.5)	+.00024 (+0.8)	-112 (-1.3)	+71.6 (+1.7)		-34.3 (-2.9)		-3.86 (-1.3)				.881
(1-6)	-.00024 (-0.2)	+.00016 (+0.6)	-142 (-2.1)	+65.5 (+2.1)		-46.0 (-5.2)		-5.63 (-2.6)	+.874 (+4.3)	+.215 (+3.8)	-4.41 (-5.0)	.944
(1-7)	-.00090 (-0.4)	+.00037 (+1.3)	-91.3 (-1.2)	+58.0 (+1.6)		-51.2 (-5.2)		-4.63 (-1.9)		+.242 (+3.9)	-5.27 (-5.4)	.928

^a Numbers in parentheses are *t*-ratios.^b MEINC2 is mean family income of the second lowest quartile of families which receive income.^c MEINC4 is mean income of highest quartile of families.^d UNEMMC is male civilian unemployment rate.^e MOBILR is proportion of population over five years of age that lived elsewhere five years previous to the 1960 census.^f NWHITE is proportion of population that is nonwhite.^g FORSTK is percentage of persons of foreign stock in the community.^h SPDVFM is proportion of females over 14 years of age who are separated or divorced.ⁱ FAMILY is proportion of children living with both parents.^j DILAP is proportion of dwellings classified as dilapidated.^k DENSE is dwelling units per net residential acre.^l MSCOOL is median years of schooling of adult population.

SPDVFM is not significant. Further, two economic variables, *MEINC4*, and *UNEMMC*, cease to be significant when *FAMILY* is substituted for *SPDVFM*; there is now only one economic variable, *MEINC2*, which exerts a significant effect on delinquency.

The inclusion of alternative and additional taste variables further reduces the effect of income on delinquency. Moreover, these variables are a priori as reasonable as those used by Fleisher. For example, Fleisher uses *NWHITE* to measure taste factors associated with race, finding that it "... adds almost nothing to the explanatory power of the regression model..." (Fleisher, Mar. 1966, p. 133). However, an alternative but analogous taste variable may be used which does add to the model's explanatory power. This variable is the percentage of first- and second-generation Americans—persons of foreign stock—in the community (*FORSTK*). In equations (1-3) and (1-4), *FORSTK* is added to (1-1) and (1-2), respectively, and is significant in each, with a negative sign. Comparison of these respective pairs of equations also shows that *UNEMMC* becomes more strongly negative when *FORSTK* is included, and that *MOBILR* ceases to be significant. *MOBILR* is often used as a measure of the concept of "anomie," or social rootlessness, as pointed out by Fleisher. *FORSTK* may also represent this concept; persons of foreign stock tend to cluster in American cities, and they frequently seek to preserve the culture of their place of origin. The regression results indicate that these variables may be measuring similar phenomena, to some extent.

There are also interesting differences between equations (1-3) and (1-4). *MEINC4* becomes insignificant when *FAMILY* replaces *SPDVFM*. *NWHITE* has a positive insignificant coefficient in (1-3); in (1-4) it has a negative coefficient which is barely significant, in Fleisher's sense.² Finally,

² The negative sign for *NWHITE* in equation (1-4) is perhaps unexpected; however, it may be consistent with sociological studies of urban family life. Further reference will be made to this point after the analysis of delinquency by subgroups.

FAMILY again has the expected sign, but *SPDVFM* is negative and is also significant in Fleisher's sense, leading to the unlikely conclusion that children from broken families are less prone to delinquency.

Equation (1-5) is similar to (1-4) with *NWHITE* removed; the other variables are not greatly affected.

Equation (1-6) introduces three additional variables which have often been regarded as causes of delinquency. Two of these measure housing conditions: the proportion of dwelling units classified as "dilapidated" in the 1960 Census of Housing (*DILAP*); and the number of dwelling units per net residential acre (*DENSE*). The third is the median years of schooling of the adult population (*MSCOOL*).

Two of these variables are theoretically closely related to income. Housing quality is of course to some extent a function of income; persons living in dilapidated housing may very likely be doing so because they cannot afford housing of standard quality. However, it has often been argued that "slums" themselves exert an independent effect on the tendency to commit delinquent acts.³

MSCOOL is also closely related to income, but the causal relationship is more likely to run in the opposite direction: a higher level of schooling as a child is likely to lead to a higher level of income as an adult. It is also likely to change the tastes of the individual, including his attitudes toward crime and delinquency; when he becomes a parent, the attitudes of his children may also be changed. This effect of schooling on delinquency could occur whether or not his income is increased. A model of delinquency which ignores schooling may be misspecified; income may be to some extent a proxy for schooling.⁴

³ This line of argument has frequently been employed as partial justification of public housing and urban renewal. (See Gist and Halbert, pp. 465-66; Ratcliff, pp. 432-33; Rothenberg, pp. 307-12.) Traces of this hypothesis may even be found in Marshall (p. 2).

⁴ In his 1966 study, *The Economics of Delinquency*, Fleisher employs *MSCOOL* as an alternative taste vari-

Equation (1-6) demonstrates the importance of the taste variables in determining the effect of income. All six taste variables are significant in (1-6), with reasonable signs; only one economic variable, *UNEMMC*, is significant, but with the "wrong" sign. Both measures of income are now insignificant.

The income and unemployment variables are highly collinear with the variables representing tastes. At the two-tail, one percent level, *MEINC2* and *UNEMMC* are significantly correlated with each taste variable, and *MEINC4* is correlated with all except *MOBILR*, *FORSTK*, and *DENSE*. Fleisher has discussed the problem of multicollinearity and has apparently sought to deal with it by restricting the number of variables used to represent tastes; he argues that "it is important . . . to beware of inadvertently including variables that represent the same economic factor twice in the same regression" (Fleisher, 1966, p. 33). This is undoubtedly true, but there is little if any theoretical basis for arguing that the additional taste variables do in fact represent economic factors such as income, rather than tastes. (Fleisher's argument is perhaps strongest with respect to *DILAP*, but even if *DILAP* is excluded on the grounds that it is largely a function of income, as in equation (1-7), the significance of income is negligible.)

Fleisher also analyzed delinquency within Chicago on the basis of subgroups of communities, stratified by *SPDVFM*. He felt that this stratification improved the specification of the model (p. 131). In this comment, communities will be stratified on the basis of *FAMILY*, since it is preferable to *SPDVFM* as a measure of the incidence of children from broken families. Two subgroups are analyzed in Tables 2 and 3: the "Low" sub-group of Table 2 consists of 22

communities in which fewer than 86.0 percent of the children live with both parents; the "High" subgroup contains the remaining 52 communities. This uneven division reflects the negatively skewed distribution of *FAMILY*; 86 percent is a fairly clear line of demarcation between the lower tail and the main body of the distribution, while there is no similarly clearcut dividing line at any higher value of *FAMILY*.⁵

The regression results for the low subgroup are similar to those for all 74 communities. The income variables are insignificant, particularly when the additional taste variables are included; *UNEMMC* is significant in Fleisher's sense, but again with the wrong sign; the taste variables again have the expected signs and (with the exception of *DENSE*) are significant, or nearly so.

By contrast, the income and unemployment variables are generally significant for the high subgroup, shown in Table 3, and they have the "right" signs. Among the taste variables, only *FORSTK* is significant in the usual sense, although *NWHITE*, *DENSE*, and *MSCOOL* are significant in Fleisher's sense. As in the full set of 74 communities, the exclusion of *DILAP* in equations (2-3) and (3-3) has no effect on the economic variables.

II. Possible Implications for Policy

In evaluating the possible implications of these three sets of results, it is useful to look first at the differences between the subgroups. In particular, the incidence of delinquency is much higher in the low subgroup: it has a delinquency rate of 67.6 per 1000 male juveniles (between 12 and 16 years of age), while the high subgroup rate is 17.3. Put another way, the communities in the low subgroup contain 59 percent of the delinquents, but only 32 percent of the juveniles.⁶

able to *SPDVFM* (pp. 35-36; 92-97). The effects on income are somewhat similar to those shown in this note although they are less pronounced. It is not clear, however, why *MSCOOL* should be regarded as an *alternative* to *SPDVFM* (or *FAMILY*); it seems rather to be an additional variable measuring an additional aspect of tastes.

⁵ This procedure differs from Fleisher's; he used three subgroups, having "High," "Medium," and "Low" *SPDVFM*. The Low *FAMILY* and High *SPDVFM* subgroups are very nearly identical, having 21 communities in common. Four others are included only in the *SPDVFM* subgroup, and one is only in the *FAMILY* subgroup.

⁶ The means of the delinquency rates for the two sub-

TABLE 2—LOW FAMILY SUBGROUP (22 COMMUNITIES)

	MEINC2	MEINC4	UNEMMC	MOBILR	NWHITE	FORSTK	FAMILY	DILAP	DENSE	MSCOOL	r ²
(2-1)	-.0101 (-1.6)	-.00099 (-1.4)	-607 (-2.7)	+23.3 (+0.9)	-74.9 (-1.9)	-207.1 (-2.4)	-15.90 (-1.9)				.784
(2-2)	+.00046 (+0.1)	-.00032 (-0.4)	-301 (-1.6)	+45.1 (+2.5)	-24.8 (-0.9)	-99.5 (-1.8)	-10.91 (-1.8)	+1.20 (+4.0)	+.101 (+0.8)	-6.23 (-2.1)	.944
(2-3)	+.00027 (+0.4)	+.00093 (+0.8)	-109 (-0.4)	+31.7 (+1.2)	-24.0 (-0.6)	-99.3 (-1.2)	-8.33 (-0.9)		+.209 (+1.1)	-9.66 (-2.2)	.862

For definitions of variables, see Table 1 fnn.

TABLE 3—HIGH FAMILY SUBGROUP (52 COMMUNITIES)

	MEINC2	MEINC4	UNEMMC	MOBILR	NWHITE	FORSTK	FAMILY	DILAP	DENSE	MSCOOL	r ²
(3-1)	-.0094 (-4.5)	+.00086 (+3.4)	+164 (+2.4)	-0.208 (-0.02)	-26.0 (-1.96)	-36.1 (-4.4)	+0.41 (+0.1)				.819
(3-2)	-.0057 (-1.9)	+.00067 (+2.1)	+143 (+2.0)	+4.57 (+0.4)	-18.3 (-1.3)	-37.4 (-4.4)	-1.65 (-0.4)	-.0310 (-0.1)	+.0892 (+1.5)	-1.62 (-1.8)	.837
(3-3)	-.0056 (-2.0)	+.00066 (+2.2)	+142 (+2.0)	+4.75 (+0.5)	-18.2 (-1.4)	-37.2 (-4.5)	-1.77 (-0.5)		+.0898 (+1.6)	-1.62 (-1.8)	.837

For definitions of variables, see Table 1 fnn.

This difference in delinquency rates between subgroups implies that economic factors are significantly related to delinquency only where delinquency itself is relatively unimportant. This in turn implies that changes in these economic factors will have little effect on delinquency. For example, an increase of \$500 per year in family income would change the delinquency rate by $(-.0057 + .00067)500 = 2.52$ per 1000 juveniles, in the high subgroup. This represents 14.6 percent of the high subgroup's delinquency, but it is only 5.9 percent of the delinquency in the city of Chicago overall. By contrast, according to Fleisher's regression results, such an increase in income would reduce the delinquency rate in Chicago by 20 percent.⁷ Further, increases in family income within the poorest part of the city would have no effect on delinquency.⁸

It is also true that reducing unemployment will reduce delinquency in the high subgroup where there is relatively little of either. The mean of *UNEMMC* in this subgroup is 3.3 percent; reducing it to zero will reduce delinquency by 4.75 per 1000, which is 27.5 percent of the high subgroup's delinquency, and 11.2 percent of the city's. Even for this subgroup, however, the elasticity of delinquency with respect to unemployment is .05, which is lower than any of Fleisher's elasticity estimates for the Chicago communities (p. 130).

For the low subgroup, and for the city as a whole, a reduction in *UNEMMC* has the effect of raising the delinquency rate, but the coefficients are generally not significant. Although his results did not show a similar pattern, Fleisher has suggested an explanation for it (p. 136). In the low subgroup, un-

employment is likely to be a normal phenomenon; it probably does not represent a deviation of current income from normal income to the same extent as it does in the high subgroup. Children in the low subgroup may not view parental unemployment as representing a narrowing of their own probable future opportunities to engage in legitimate activity. By the time they are old enough to appear in the delinquency statistics, they may have come to expect their parents to be unemployed relatively frequently; whether or not their parents are actually unemployed at any given time may have no effect on their behavior. If this hypothesis is correct, it suggests that policies to combat unemployment will also combat delinquency only if they succeed in changing children's expectations about the probability of unemployment; that is, only if they are successful in reducing unemployment for long periods. Should the low unemployment rate of the late 1960's continue until the 1970 Census, this hypothesis predicts that coefficients for *UNEMMC* will be similar in the two subgroups at that date.

The results for some of the taste variables may have policy implications, although the existence of multicollinearity indicates that these implications should be treated as more in the nature of tentative suggestions. For example, the significance of the housing variables suggests that such policies as public housing and urban renewal may be more effective in fighting delinquency than policies designed to raise income or reduce unemployment, to the extent that the housing policies are successful in reaching their goals of eliminating slums and reducing density.⁹ Urban renewal projects in particular usually aim to reduce the density of population in the project area; however, the coefficients of *DENSE* in (2-2) and (2-3) indicate that reducing population density is not likely to

groups are significantly different; stratifying by *FAMILY* is almost the same thing as stratifying by delinquency. The means of the two subgroups also differ significantly for all of the independent variables except *MISC00L*.

⁷ The calculations are based on the coefficients for *MEINC2* and *MEINC4* in equations (3-2) and (1-1), respectively. The income figure of \$500 is used in order to facilitate comparison with Fleisher's calculations (p. 134).

⁸ All communities having *MEINC2* less than \$5,000 lie in the low subgroup.

⁹ *DILAP* is also significantly positively correlated with delinquency, which may be interpreted to mean that the tastes for substandard housing and delinquency may both be derived from some underlying phenomena. If this is true, then the housing policies cited may not generate a reduction in delinquency, unless they also affect these underlying phenomena.

reduce delinquency in the most delinquency-prone communities.

MSCOOL, strictly speaking, refers to the schooling of adults rather than juveniles. It would probably be preferable to use a measure of the schooling of the present generation of juveniles, such as a dropout rate. Unfortunately, data are not available to calculate dropout rates by community within Chicago, although Fleisher's time-series analysis of delinquency (1966, pp. 78-84) strongly indicates that the dropout rate is highly correlated with the delinquency rate. If *MSCOOL* is regarded as a proxy for the dropout rate, the regression results suggest that programs which encourage juveniles to undertake additional schooling, such as Operation Head Start, are more likely to reduce delinquency than are anti-poverty programs in general.

The regression results for *NWHITE*, particularly in the subgroups, imply that, *ceteris paribus*, Negroes are less likely to be delinquent than whites;¹⁰ nearly all coefficients for *NWHITE* in Tables 2 and 3 are significant in Fleisher's sense. Perhaps most economists, and almost certainly most laymen, would find this result unlikely. The explanation may lie in the relationship between *NWHITE* and *FAMILY*. Sociologists have noted a high incidence of instability among lower-class Negro families, dating back to the period of slavery and continuing until the present;¹¹ some have suggested that this pattern of broken families tends to increase delinquency among Negroes because it leaves the children unsupervised while the mother has to work (Gist and Halbert, pp. 414-15). Broken families are more normal for Negroes than whites; in the Chicago data, 38 percent of nonwhite children under 18 did not live with both parents in 1960, compared to only 10 percent of white children. Since a broken family is more of a departure from community norms for a white family, it may be more likely to generate

socially unacceptable behavior among whites, including delinquency. For a given level of *FAMILY*, in other words, whites are more likely to be delinquent, since their usual pattern of life has been disrupted to a greater extent. The regression results are consistent with this hypothesis; they indicate that the observed higher incidence of delinquency among Negroes is in fact generated by the higher incidence of broken families and unsupervised children.

III. Conclusion

In general, the regression results show a different pattern from that shown by Fleisher. Traditional sociological "taste" variables, such as "anomie" and "the absence of a strong father-figure in the home" have significant effects on delinquency; economic variables appear to exert no effect, although the evidence for *UNEMMC* suggests that changes in permanent income may exert some effect on delinquency in the long run.¹²

This analysis has been based on only one of several sets of data utilized by Fleisher. However, as pointed out earlier, Fleisher regarded this data as probably the most useful for the study of delinquency (perhaps along with data for Chicago suburbs), for several reasons, such as higher variability of social and economic factors within cities than between them, and relative homogeneity of the reporting agencies. Furthermore, in his regressions the economic variables generally had higher coefficients and *t*-ratios in the Chicago communities than in the other sets of data. In particular, both *MEINC2* and *UNEMMC* are significant, in Fleisher's sense, only in three of the eleven non-Chicago regressions which also include his measure of family structure (*SPDVFM*);

¹² It is perhaps worth pointing out that these results do *not* imply that antipoverty or other programs designed to raise incomes are either economically inefficient or normatively undesirable; all the results imply is that such programs are not likely to generate additional benefits to society in the form of lower delinquency rates, at least in the short run. Since Fleisher has always been interested in the policy implications of his results, the lack of such benefits is worth stressing. (See Fleisher 1963, p. 543, and Mar. 1966, pp. 136-37.)

¹⁰ In Chicago, 97 percent of nonwhites are Negroes Kitagawa and Taeuberg (p. 246).

¹¹ The correlation between *NWHITE* and *FAMILY* is extremely high in Chicago (-.94 for the full set of 74 communities).

MEINC4 is significant in only one of nine. When combined with the results of the present analysis, these results provide no basis for concluding that there is any "effect of income on delinquency."

REFERENCES

- B. M. Fleisher, "The Effect of Unemployment on Juvenile Delinquency," *J. Polit. Econ.*, Dec. 1963, 71, 543-55.
- , "The Effect of Income on Delinquency," *Amer. Econ. Rev.*, Mar 1966, 56, 118-37.
- , *The Economics of Delinquency*. Chicago 1966.
- N. P. Gist and L. A. Halbert, *Urban Society*, 4th ed. New York 1956.
- E. M. Kitagawa and K. E. Taeuber eds., *Local Community Fact Book Chicago Metropolitan Area 1960*. Chicago 1963.
- A. Marshall, *Principles of Economics*, 8th ed. New York 1948.
- R. U. Ratcliff, *Urban Land Economics*. New York 1949.
- J. Rothenberg, "Urban Renewal Benefits," in R. Dorfman, ed., *Measuring Benefits of Government Investments*, Washington 1965.

The Effect of Income on Delinquency: Reply

By BELTON M. FLEISHER*

John Weicher's comment emphasizes the inherent difficulty of empirical investigation when the underlying behavioral model is only vaguely specified. Not only is there lack of agreement on what are the important variables theoretically influencing delinquency, but even for those variables generally agreed on in theory, there is little we have to go on if we want to identify their observable counterparts. Thus, it is difficult for Weicher to make a case that conclusions drawn from one study "... are vitiated" by the results of another study which has simply thrown more variables into the statistical hopper.

I cannot help but agree with Weicher that the variable *FAMILY* has at least as good a claim to represent "tastes" for delinquency as the variable *SPDVFM*, which I used. Thus, I find the results of equation (1-2) at least as satisfactory a representation of the actual forces at work on delinquency as those of equation (1-1). Unfortunately, such a comparison is not possible in Tables 2 and 3, as the counterparts of equation (1-2) in these tables are not shown. However, beyond this, I find it impossible to agree with Weicher that his results represent an improvement on my own. He introduces the variable *FORSTK* suggesting that "anomie" may be represented by this concept; if it is, then it would appear that its regression co-

efficient should be positive in sign—not negative as in equations (1-3) and (1-4). Thus, one is left in the position of not knowing whether to reject the entire equation because the "anomie" variable has not been properly represented, or whether to accept it, changing one's maintained hypothesis about the amount of "anomie" in foreign-born or second-generation neighborhoods. Without further justification by Weicher, I find it impossible to draw any conclusion about the effects of income and unemployment on delinquency from the equations containing *FORSTK*, for the reason cited.

For similar reasons, I cannot draw any conclusions from the regressions resulting from the addition of the variables added in equations (1-6) and (1-7). I am aware that incorrect specification of the delinquency model may have caused me to draw wrong inferences about underlying relationships and the effects of policies altering the levels of income and employment on juvenile delinquency. It was my belief when reporting the results of my study that the absence of a well-thought-out theory of delinquency makes empirical work in the field difficult and policy prescriptions risky; I still hold that belief today. I do not think that adding new variables whose relationships to delinquency and its causes are specified only in terms of "may" and concepts similarly difficult to subject to empirical test can increase our knowledge of the causes and cures of juvenile delinquency.

* The author is professor of economics at The Ohio State University.

ERRATA

1. There are errors in the notation of Lester Telser's discussion on page 123 of the May 1969 *Papers and Proceedings* issue. The three lines starting with equation (3) should read

$$(3) \quad r - \bar{r} = (A/A - \bar{r})w_A$$

where

$$w_A = A/(A + K), \quad A = \frac{dA}{dt}$$

and

$$\dot{A} - a = Ad_A$$

2. D. Levhari and Don Patinkin, "The Role of Money in a Simple Growth Model," this *Review*, Sept. 1968, 58, p. 720, Lines 11 and 12 should read

"Thus the real balance effect implicit in (8) is

$$\partial C_F / \partial (M/p) = (1 - s)(\mu + r)."$$

The error was pointed out by Robert J. Moreland of the Government of the Province of New Brunswick, Treasury Board, Fredericton, N. B., and by J. K. Stephens and T. G. Malliaris, University of Oklahoma.



NOTES

The 1970 Summer Institute in Behavioral Science and Law will be held on the Madison campus of the University of Wisconsin. Graduate students in the social sciences and law students who have completed a year of legal studies are eligible for admission. For information on courses, write to Professor Joel B. Grossman, director, Summer Institute in Behavioral Science and Law, Social Science Building, University of Wisconsin, Madison, Wisconsin 53706.

The Survey Research Center of the University of Michigan will hold a Summer Institute in Survey Research Techniques for the twenty-third consecutive year in 1970. There will be two four-week sessions: The first from June 30 to July 24 and the second from July 28 to August 21. They may be taken independently or successively.

The Institute is designed to meet some of the educational and training needs of men and women engaged in business and governmental research and other statistical work, and graduate students and university instructors interested in quantitative research in the social sciences.

For further information, write to Mrs. Betty F. Jennings, Senior Administrative Assistant, Office of the Director, Survey Research Center, The University of Michigan, P.O. Box 1248, Ann Arbor, Michigan 48106.

New Journal: *Social Theory and Practice*

This journal is intended to provide a forum for the expression of important and controversial social and political issues. The editors wish to reestablish the relevance of philosophical criticism and a philosophic overview to the solution of pressing social problems. The journal will publish original work in social philosophy by authors from all relevant disciplines, including the humanities, the social sciences and the natural sciences, as well as those engaged in practical social and political action. Both constructive and critical work at all levels and on all social issues are welcome.

A part of each issue will be devoted to a specific theme. The suggested themes for the first three issues are the following (the deadline for manuscripts is added in parentheses):

- 1) The Relationship between Social Theory and Practice (December 1, 1969).
- 2) Class Role and Class Interest (July 1, 1970).
- 3) International Economic Realities and Ideologies (December 1, 1970).

All communications, manuscripts, subscriptions, and books for review should be sent to the Editor, K. T. Fann, Department of Philosophy, The Florida State University, Tallahassee, Florida 32306.

The National Science Foundation is seeking grant proposals for research and education in the area of science planning and policy. The grants which are for institutionwide, multidisciplinary programs are designed to: (a) meet the urgent need to understand and solve the many complex science policy issues which confront the nation; (b) stimulate academic institutions to use their intellectual resources to analyze, on an interdisciplinary basis, important issues in science policy and planning; (c) meet the need for a more adequate supply of educated citizens and experts in this field; (d) bring a wider range of informed viewpoints to bear on international, national, state, regional, and local science policy issues.

Institutions currently receiving grants under this program include: Harvard, Massachusetts Institute of Technology, University of Virginia, University of Illinois, State University of New York at Albany, Cornell, and the University of Indiana.

These grants help to support research and special seminars on science planning and policy related to such problems as: the use of science in international affairs; scientific and technical manpower; environmental management; technology and the city; nuclear energy, the law and international affairs; the effects of new educational technology; effects of technology on economic growth; the organization of large-scale technological projects.

For further information write the Office of Planning and Policy Studies, National Science Foundation, Washington, D.C. 20550.

The Asia Foundation has provided the American Economic Association with a fund to be used to assist students and visiting scholars from Asia to attend the annual meeting of the American Economic Association. The maximum amount available to an individual applicant is \$100. Inquires should be addressed to the AEA, 629 Noyes Street, Evanston, Illinois 60201.

The Joint Council on Economic Education, with the advice of the AEA Committee on Economic Education, is conducting a contest for multiple choice questions requiring application of economic principles to realistic situations. First prize is \$1,000, second prize \$500, third prize \$250, and additional prizes of \$25 per usable question submitted. Each contestant must submit at least five questions, which will become the property of the JCEE. Deadline: September 15, 1970. For further details, write Dr. Arthur L. Welsh, Joint Council on Economic Education, 1212 Avenue of the Americas, New York 10036.

The Committee on International Exchange of Persons has recently issued the Directory of Visiting Scholars in the United States awarded grants under the Ful-

bright-Hays Act for 1969-70. The Committee also expects to issue in March 1970 a list of foreign scholars available under the Fulbright-Hays program for appointments in American colleges and universities for the academic year 1970-71. The list to be issued in the spring will include information about scholars nominated by the binational Educational Commissions and Foundations abroad for Fulbright-Hays travel grants covering costs of round-trip transportation from the home country to the United States, provided arrangements can be completed for lecturing or research appointments with appropriate stipends at American institutions of higher learning.

Persons interested in receiving a copy of the current Directory for the academic year 1969-70 or a copy of the forthcoming list of candidates to be issued in March 1970 may write to: Miss Grace E. L. Haskins, Program Officer, Committee on International Exchange of Persons, 2101 Constitution Avenue, N.W., Washington, D.C. 20418 (or telephone 202-961-1648).

Summer *NSF* Institute in management science and operations research for college professors in management science, operations research, applied mathematics, economics, health and hospital administration to be held June 15 through July 17, 1970 at the University of Colorado, Boulder. Participants receive stipend and travel expenses.

For further information, correspond with: Dr. Donald R. Plane, Co-Director, Summer Institute, Division of Management Science, Business Building, University of Colorado, Boulder, Colorado 80302.

Stanford University will sponsor a Summer Institute in Urban Economics for college and university teachers of economics, June 28-July 25, 1970. The program is supported by the National Science Foundation.

Twenty-five participants will be selected. The purposes of the Institute are to train the participants in the subject matter and methodology of urban economics, and to assist them to build courses of instruction in urban economics at their own institutions. For application forms and more information, write: Professor Henry M. Levin, *NSF* Institute in Urban Economics, School of Education, Stanford University, Stanford, California 94305.

A special summer program at Massachusetts Institute of Technology, August 10-21, 1970, is designed to present today's urban problems and develop the economic methodology needed to analyze and assess both the problems and their solutions in the following area: urban renewal, discrimination, state and local public finance, externalities, zoning, transportation, location and land use, poverty, crime and pollution. For further information, write: Director of the Summer Session, E19-356, M.I.T., Cambridge, Massachusetts 02139.

Deaths

Paul Homan, *Resources for the Future*, Washington, July 3, 1969.

Donald L. Richard, department of accounting, University of Kentucky, Feb. 15, 1969.

Robert J. Tompkin, Ohio State University, Nov. 19, 1969.

Retirements

Friedrich Baerwald, professor of economics, Fordham University, June 1970.

Richard C. Creek, associate professor of economics, Colorado State University, June 30, 1970.

Lewis K. Johnson, professor of administration, School of Commerce, Economics, and Politics, Washington and Lee University, June 1, 1969.

Howard M. Teaf, Jr., professor emeritus, Haverford College.

Lawrence W. Towle, Trinity College, Jan. 1970.

Visiting Foreign Scholars

George L. Beckford, University of the West Indies, Kingston, Jamaica: Food Research Institute, Stanford University, 1969-70.

Pierre M. Candau, Université d'Aix-Marseille: Graduate School of Business Administration, University of California, Los Angeles, 1969-70.

Emil Claassen, National Center of Scientific Research: visiting professor of economics, University of Chicago, winter and spring 1970.

Vassil Dimitrov, Economic Institute of Bulgarian Academy of Sciences: visiting professor of economics, Florida State University, fall 1969.

S. Herbert Frankel, Oxford University: visiting professor of economics, University of Virginia, second semester, 1969-70.

Seiko Higa, Kobe University of Commerce: visiting assistant professor of economics, Purdue University.

K. Jayachandran, Government of India Planning Commission: Graduate School of Business Administration, University of California, Los Angeles, 1969-70.

Murray C. Kemp: Ford Foundation research professor of economics, University of California, Berkeley, 1969-70.

Kenneth A. Leslie, University of the West Indies, Kingston, Jamaica: Food Research Institute, Stanford University, 1969-70.

William Mackenzie, McMaster University, Ontario: Food Research Institute, Stanford University, 1969-70.

R. Eugene Moran, Ateneo de Manila University, The Philippines: Graduate School of Business Administration, University of California, Los Angeles, 1969-70.

Glen A. Mumey, University of Saskatchewan: visiting professor of finance, Purdue University, 1969-70.

Jose R. Nacu, La Salette College, The Philippines: Graduate School of Business Administration, University of California, Los Angeles, fall and winter, 1969-70.

Richard S. Odingo, University College, Nairobi: Food Research Institute, Stanford University, until December 1970.

Premysl Pergler, Czechoslovakia: visiting lecturer,

Graduate School of Business Administration, University of California, Los Angeles, 1969-70.

Umberto Sulpasso, University of Bari, Italy: Graduate School of Business Administration, University of California, Los Angeles, 1969-70.

Satori Takayanagi, Chuo University, Tokyo: Graduate School of Business Administration, University of California, Los Angeles, 1969-70.

Jean-Marie Toulouse, University of Montreal: Graduate School of Business Administration, University of California, Los Angeles, 1969-70.

Malcolm C. Urquhart, Queen's University, Kingston: University of California, Berkeley, 1969-70.

Hartmut Wächter, University of Göttingen, Germany: Graduate School of Business Administration, University of California, Los Angeles, fall and winter, 1969-70.

Promotions

John W. Allen: associate professor of economics, Texas A&M University.

James Anderson: assistant professor of economics, Boston College.

J. Fred Bateman: associate professor of business economics and public policy, Indiana University.

Charles P. Blitch: professor of economics, Old Dominion University, fall 1969.

Earl W. Bracey: associate professor of business education, Old Dominion University, fall 1969.

Gerald E. Breger: professor of economics, University of South Carolina.

Robert J. Carlsson: associate professor of economics, University of South Carolina.

Wayne Cole: associate professor, School of Business Administration, East Texas State University.

B. Joe Colwell: associate professor of economics, University of Texas, Austin.

Anne S. Daughtrey: professor of business education, Old Dominion University, fall 1969.

Charles M. DeHority: associate professor of business management, Old Dominion University, fall 1969.

Mary M. Fish: professor of economics, University of Alabama.

William D. Gunther: assistant professor of economics, University of Alabama.

Peter Gutmann: professor of economics, Bernard M. Baruch College.

James G. Hilton: professor of economics, University of South Carolina.

Yutaka Horiba: assistant professor of economics, Macalester College.

Robert L. Jack: associate professor, agricultural economics, West Virginia University.

James D. Knowles: assistant professor of economics, Dartmouth College.

William C. Perkins: associate professor of quantitative business analysis, Indiana University.

Donald D. Rohdy: professor of economics, Colorado State University.

Walter F. Rohrs: assistant professor, Wagner College.

Thomas J. Rothenberg: associate professor of economics, University of California, Berkeley.

A. Allan Schmid: professor of agricultural economics and resource development, Michigan State University.

William L. Silber: associate professor of economics, Graduate School of Business Administration, New York University.

A. M. Sinclair: professor of economics, Dalhousie University.

P. A. V. B. Swamy: associate professor of economics, State University of New York, Buffalo, Sept. 1969.

John F. Vallery: associate professor of economics, University of Alabama.

Robert L. Winkler: associate professor of quantitative business analysis, Indiana University.

Administrative Appointments

William M. Baird: chairman, department of economics, The College of Wooster.

George F. Break: chairman, department of economics, University of California, Berkeley.

Robert C. Burton: director, Center for Metropolitan Studies, Old Dominion University, fall 1969.

Robert S. Carlson: dean, Babcock School of Business, Wake Forest University, fall 1969.

Donald D. Christenson: associate dean, College of Business Administration and director, graduate programs in business, Wichita State University.

Thomas R. DeGregori: acting chairman, department of economics, University of Houston, 1969-70.

John F. DeVogt: chairman, department of administration, School of Commerce, Economics, and Politics, Washington and Lee University, June 1969.

Robert B. Ekelund, Jr.: director of undergraduate programs, department of economics, Texas A&M University.

Gladys Ellenbogen: chairman, department of economics, Montclair State College, Sept. 1969.

Frederick J. Glover: chief executive officer, International Trade Centre, Unctad/Gatt, Geneva, Switzerland.

Charles M. Hewitt: dean, junior division, Indiana University.

Ole S. Johnson: dean, School of Business Administration, Old Dominion University, summer 1969.

Paul E. Junk: chairman, department of economics, University of Missouri-Columbia.

Archie Kleingartner: associate dean and chairman, department of business administration, Graduate School of Business Administration, University of California, Los Angeles.

J. David Lages: acting head, department of sociology and economics, Southwest Missouri State College.

William McWhinney: assistant dean for planning, Graduate School of Business Administration, University of California, Los Angeles.

Morris L. Mayer: chairman, department of marketing, College of Commerce and Business Administration, University of Alabama.

Sid Mittra: associate professor, chairman, department of economics, Oakland University, July 1969.

Thomas D. Moore: professor of economics, director, Center for Economic Education, University of Alabama.

Kenneth C. Nobe: professor, chairman, department of economics, Colorado State University, Dec. 1969.

Jeffrey B. Nugent: acting chairman, department of economics, University of Southern California, Los Angeles, 1969-70.

Harold C. Passer: Assistant Secretary of Commerce for Economic Affairs, Washington.

Tom L. Powrie: chairman, department of economics, University of Alberta, 1969-70.

William T. Reece: chairman, professor of accounting, Old Dominion University, fall 1969.

John K. Ryans, University of Kentucky: assistant dean, Kent State University.

John M. Scheidell: chairman, economics department, Florida Atlantic University, 1969-70.

Mark B. Schupack: chairman, department of economics, Brown University.

Lester Stone: assistant to the dean, School of Business Administration, East Texas State University.

Ronald L. Stump: acting dean of Appalachian Center; acting director, Cooperative Extension Service, West Virginia University.

Peter Vaill: vice chairman, department of business administration, Graduate School of Business Administration, University of California, Los Angeles.

Michael Y. Yoshino: vice chairman, department of business administration, Graduate School of Business Administration, University of California, Los Angeles.

New Appointments

James G. Agee: instructor, Auburn University, 1970.

Klaus F. Alt: research associate, department of economics, Iowa State University.

Bernard E. Anderson: assistant professor of industry, Wharton School of Finance and Commerce, University of Pennsylvania.

Lloyd C. Atkinson: assistant professor of economics, University of Maryland, fall 1969.

Gary S. Becker, Columbia University: Ford Foundation visiting research professor of economics, University of Chicago, 1969-70.

John J. Cardwell: professorial lecturer, Graduate School of Business, University of Chicago, 1969-70.

William J. Chambers: instructor in economics, The College of Wooster.

Kenneth W. Clarkson: acting assistant professor of economics, University of Virginia.

Susan Cochrane: assistant professor of economics, University of South Carolina.

John E. Cooper, University of Idaho: assistant professor of economics, Lewis and Clark College, fall 1969.

Edward M. Cross: associate professor of business management, Old Dominion University, fall 1969.

Thomas E. Daves, University of Minnesota: research associate, Agency for International Development, Tunisia, 1970-72.

David Denny: lecturer in economics, University of Maryland, fall 1969.

Johannes A. Dockel: research associate, department of economics, Iowa State University.

Steven W. Dobson: instructor of economics, Dartmouth College.

K. William Easter, University of Minnesota: associate professor of agricultural economics, The Ford Foundation, New Delhi.

Wendell Edwards, North Texas State University: assistant professor of accounting, School of Business Administration, East Texas State University.

Michael Etzel, Colorado University: assistant professor, University of Kentucky.

Ruth A. Fabricant, Federal Reserve Board: senior research analyst, Office of Comprehensive Planning, New York City Department of City Planning.

Richard J. Fleming: associate professor of business management, Old Dominion University, fall 1969.

Richard B. Freeman, Harvard University: assistant professor, University of Chicago, Jan. 1970.

William P. Glade, Jr., University of Wisconsin, Madison: department of economics, University of Texas, Austin, spring 1970.

C. Ray Gullett, Louisiana State University: assistant professor of management, School of Business Administration, East Texas State University.

Herbert H. Hand, The Pennsylvania State University: assistant professor of business administration, Indiana University, Sept. 1969.

Bennett Harrison: lecturer in economics, University of Maryland, fall 1969.

George W. Hettenhouse: assistant professor of finance, Graduate School of Business, Indiana University, fall 1969.

Neal Higgins, University of Nebraska: assistant professor of economics, School of Business Administration, East Texas State University.

Lanetta Holiman, North Texas State University: instructor in accounting, School of Business Administration, East Texas State University.

Herbert A. Howard, Indiana University: assistant professor, Auburn University, Mar. 1969.

Keith Johnson, University of Illinois: assistant professor, University of Kentucky.

Charles P. Jones: assistant professor of finance, Graduate School of Business, Indiana University, fall 1969.

James A. Kane: instructor, department of economics, University of Massachusetts.

Harvey B. Karp: assistant professor of business management, Old Dominion University, fall 1969.

Paul C. Kettler: instructor, Graduate School of Business, University of Chicago, 1969-70.

Abdul G. Khan: head of programme on institutional management in higher education, Center for Educational Research and Innovation, Organization for Economic Cooperation and Development, Paris.

John P. Kuehn: assistant professor of agricultural economics, West Virginia University.

John L. Lewis: instructor of economics, Southwest Missouri State College.

R. Dean Lewis, University of Arkansas: assistant professor in management, School of Business Administration, East Texas State University.

William V. Luckie, Jr.: instructor, accounting, Babcock School of Business, Wake Forest University, fall 1969.

Robert B. McBurney: instructor of economics, North Carolina State University.

Parthasaradhi Mallela: assistant professor of economics, State University of New York, Buffalo, Sept. 1969.

David Mishkin: associate professor of finance, Fordham University.

John F. Muth, Michigan State University: professor of production management, Indiana University, Sept. 1969.

Lonnie Nickles: assistant professor of economics, University of South Carolina.

Daniel Ounjian: associate professor of economics, The Fletcher School of Law and Diplomacy, Tufts University.

Lewis J. Perl: assistant professor, Cornell University, New York State School of Industrial and Labor Relations.

Edgar A. Persons: associate professor, departments of agricultural economics and agricultural education, University of Minnesota.

George J. Petrello, Montclair State College: associate professor of economics and business administration, Wagner College.

Joseph D. Phillips: associate professor of management, University of Alabama.

Joseph M. Pincus, U.S. Department of State: professor of economics, Louisiana Polytechnic Institute.

Robert Platt: assistant professor of economics, Graduate School of Business Administration, New York University.

H. Joseph Reitz, Massachusetts Institute of Technology: assistant professor, Indiana University, Feb. 1969.

Karl D. Reyer: visiting professor of marketing, Babcock School of Business, Wake Forest University, fall 1969.

Paul N. Rosenstein-Rodan, Massachusetts Institute of Technology: department of economics, University of Texas, Austin, spring 1970.

Cathy Russell, North Texas State University: instructor in business education, School of Business Administration, East Texas State University.

William Sartoris, Purdue University: assistant professor, University of Kentucky.

Theodore P. Scheinman, West Liberty State College: assistant professor of economics, Lewis and Clark College, fall 1969.

Leroy B. Schwarz: instructor, Graduate School of Business, University of Chicago, 1969-70.

Donald Shannon, University of North Carolina: assistant professor, University of Kentucky.

Caroline M. Smith: instructor of business management, Old Dominion University, fall 1969.

Kevin Sontheimer, Virginia Polytechnic Institute: assistant professor of economics, State University of New York, Buffalo, Sept. 1969.

Lynn A. Stanton: assistant professor, agricultural economics, West Virginia University.

Gabriel Temkin, University of Warsaw: lecturer, department of economics, State University of New York, Stony Brook, Feb. 1970.

Richard J. Tersine: assistant professor of business

management, Old Dominion University, fall 1969.

James M. Utterback, Massachusetts Institute of Technology: assistant professor of business administration, Indiana University, Feb. 1969.

Earl D. Wasson: assistant professor of economics, Southwest Missouri State College.

David Whitcomb: assistant professor of economics, Graduate School of Business Administration, New York University.

Ruth Wieters, Oklahoma State University: instructor in business education, School of Business Administration, East Texas State University.

Richard H. Wood: professor of business administration, University of Alabama.

Camille M. Woodward, Columbus College: instructor, Auburn University, 1969-70.

Leaves for Special Appointments

Ichak Adizes, Graduate School of Business Administration, University of California, Los Angeles: visiting lecturer, Hebrew University, Jerusalem, winter, spring and summer, 1970.

A. James Barnes, Indiana University: special assistant to Assistant Attorney General, Washington, 1969-70.

William F. Beazer, University of Virginia: visiting professor, University of East Africa, Nairobi, 1969-70.

David A. Belsley, Boston College: fellow, Center for Advanced Study, Behavioral Sciences, Stanford University.

Julian H. Bradsher, Oklahoma State University: professor-consultant, Middle Eastern Technical University, Ankara, Turkey; technical assistance program, Economic Cooperation and Development, Paris.

Leland Burns, Graduate School of Business Administration, University of California, Los Angeles: associate dean, School of Architecture and Urban Planning, University of California, Los Angeles, 1969-70.

Lehman B. Fletcher, Iowa State University: Food Research Institute, Stanford University, March-June 1970.

Gregory Grossman, University of California, Berkeley: Center for Advanced Study, Behavioral Sciences, Stanford University, 1969-70.

William T. Hogan, Fordham University: visiting professor of mineral economics, Pennsylvania State University, spring 1970.

Allen C. Kelley, University of Wisconsin, Madison: Monash University, Clayton, Australia.

Mohamad W. Khouja, Oklahoma State University: economic consultant, Kuwait Fund for Economic Development, Middle East, July 1969 to Aug. 1970.

Arnold B. Larson, University of Hawaii: visiting scholar, Food Research Institute, Stanford University, Feb.-June 1970.

Martin C. McGuire, University of Maryland: visiting research professor, York University, England, Sept. 1969-Aug. 1970.

E. W. Martin, Jr., University of Hawaii: visiting professor, Indiana University, 1969-70.

Lee R. Martin, University of Minnesota: visiting

professor of agricultural economics, University of California, Berkeley.

R. Bruce Ricks, Graduate School of Business Administration, University of California, Los Angeles: Director of the Office of Research and Home Finance and Chief Economist, Federal Home Loan Bank Board, Washington 1969-70.

Lawrence X. Tarpey, University of Kentucky: dean, School of Business, Haile Sellassie University, Ethiopia, 1969-71.

Kenneth H. Thomas, University of Minnesota: department of agricultural economics, Purdue University.

Resignations

Brenda S. Andress, Auburn University: Florida State University.

Arthur F. Belote, Old Dominion University, fall 1969.

Holland C. Blades, Auburn University.

H. Marshall Booker, Old Dominion University, summer 1969.

Charles N. Bryan, Auburn University.

Whitney Chamberlain, Old Dominion University,

summer 1969.

Robert F. Gwinner, College of Commerce and Business Administration, University of Alabama: Arizona State University, Feb. 1, 1970.

Harold M. Hochman, University of Virginia: The Urban Institute.

Lawrence D. Jones, Indiana University, June 1969.

John P. Lewis, Indiana University, March 1969.

John D. Martin, Louisiana Polytechnic Institute: Texas Technological College.

Donald P. Paterson, Auburn University.

John Rowntree, York University, June 1969.

Bikas C. Sanyal, Iowa State University: UNESCO, Paris.

Howard G. Schaller, Indiana University: Bangkok, Thailand, July 1969.

James A. Smith, Lewis and Clark College: University of Wisconsin, Sheboygan, fall 1969.

James L. Stallings, West Virginia University: Auburn University.

Charles R. Strain, Auburn University.

Abner W. Womack, Auburn University: U.S. Department of Agriculture, Washington.



EMPLOYMENT SERVICES

NATIONAL REGISTRY FOR ECONOMISTS

The National Registry for Economists was established in January, 1966, to provide a centralized nationwide clearinghouse for economists on a year-round basis. It is located in the Chicago Professional Placement Office of the Illinois State Employment Service and is staffed by experienced placement personnel, operating under the guidance and direction of Regional and National Bureau of Employment Security Professional Placement officials, and in cooperation with the American Economic Association. It is a free service. There are no registration, referral, or placement fees. Application and order forms used in the Registry are available upon request from the: National Registry for Economists, Professional Placement Center, 208 South La Salle Street, Chicago, Illinois 60604.

AMERICAN ECONOMIC ASSOCIATION VACANCIES AND APPLICATIONS

The Association is glad to render service to applicants who wish to make known their availability for positions in the field of economics and to administrative officers of colleges and universities and to others who are seeking to fill vacancies.

The officers of the Association take no responsibility for making a selection among the applicants or following up the results. The Secretary's office will merely afford a central point for clearing inquiries; and the *Review* will publish in this section a brief description of vacancies announced and of applications submitted (with necessary editorial changes). Since the Association has no other way of knowing whether or not this section is performing a real service, the Secretary would appreciate receiving notification of appointments made as a result of these announcements. It is optional with those submitting such announcements to publish name and address or to use a key number. Deadlines for the four issues of the *Review* are February 1, May 1, August 1, and November 1.

Communications should be addressed to: The Secretary, American Economic Association, 629 Noyes Street, Evanston, Illinois 60201.

Vacancies

Labor economists and/or statisticians: The measurement of the economic damages suffered by those who are wrongly injured or killed is a new area for the application of economic and statistical knowledge and research techniques. This field has been pioneered successfully and Associated Appraisers of Earning Capacity is now providing this service to attorneys on a nationwide basis. Those in it appraise the losses and are required to testify as an expert in court trials and must withstand hostile cross-examination. It is possible to engage in it as a part-time activity. Requirements: (1) minimum status of associate professor; (2) experience in research in collective bargaining agreements and fringe benefit programs; (3) location in or near a major metropolitan area. Openings now available in some 20 cities. Waiting list open in other areas. Send résumé and list of publications to: Philip Eden, Presi-

dent, Associated Appraisers of Earning Capacity, 1303 Walnut Street, Berkeley, California, 94709.

Economists, junior and senior: Quantitative research in industrial economics, urban and transportation planning, and interregional economics. Economic model building, systems analysis, and cost-effectiveness studies. Immediate positions available in suburban Washington location. For personal interview, please send résumé to: Jack G. Faucett, Jack Faucett Associates, 8605 Cameron Street, Silver Spring, Maryland, 20910.

Accounting, finance, management, marketing, international business: Faculty openings for September, 1969, upper level and graduate level; doctorate preferred; experience in teaching and business. All rank levels. Salary open—competitive—depending on qualifications. Further informa-

tion from and résumé to: Wilbur E. Benson, Dean, College of Business Administration, University of Akron, Akron, Ohio, 44304.

Marketing: Associate professor or assistant professor. Full-time career position in teaching and research, with choice of courses in general marketing, advertising, consumer behavior, or marketing research. Development of new courses is encouraged in both quantitative and behavioral aspects. Work will be mostly with undergraduates, but there are increasing needs for graduate teaching and for guidance of masters' and doctors' theses in a complete Graduate College. Salary competitive. Summer teaching optional for extra salary. Funds are available for research and related travel. College is member of AACSB and has computer facilities, short-course and institute programs. New offices and classrooms. Community is excellent for family living. Send personal data sheet, with references, to: Professor Phillip McVey, Department of Marketing, University of Nebraska, Lincoln, Nebraska, 68508.

Economic statistics, operations research, management: Three Ph.D.'s to fill the following three positions: (1) operations research (management sciences) and/or statistics and quantitative methods; (2) management and organization theory, personnel management, business policy; production management, management systems, international management. Huge M.B.A. and M.S. in C. and Ph.D. programs. Metropolitan location ideal for consulting and research. Nine hours teaching—two undergraduate courses and one graduate course. Assistant professor, \$13,500; associate professor, commensurate with experience (competitive). T.I.A.A., life, disability, and major medical insurance. Full tuition remission for children of faculty. Send résumé to: Dr. Sam Barone, Chairman, Department of Management Sciences, St. Louis University, School of Commerce and Finance, 3674 Lindell Boulevard, St. Louis, Missouri, 63108.

Economist: September, 1970. Ph.D. Strong background in economic theory to teach at both undergraduate and graduate levels. Rank and salary dependent upon qualifications. Please send résumé to: Dr. Donald Leiffer, Chairman, Department of Economics, American University in Cairo, 113 Sharia Kasr El Aini, Cairo, Egypt.

Economics: One or more openings for September, 1970, in an emerging undergraduate economics program. Most fields open. Appointment at any level but prefer Ph.D. and teaching experience. Opportunity for interdisciplinary participation if desired. Twelve-hour teaching load. Competitive salary levels. Located 125 miles southwest of Minneapolis in a community of 8,000. The College is in its third year with near-term enrollment of 4,000. Please send résumé to: Chairman, Economics Program, Southwest Minnesota State College, Marshall, Minnesota, 56258.

Business administration: Rapidly growing branch of University of Alaska, offering B.B.A. in Business Administration, B.A. in Economics, and M.B.A. in Business Administration. Requires D.B.A. or Ph.D. in business administration or dissertation stage. Nine-hour teaching load, covering graduate and undergraduate programs plus supervision of graduate students. Rapidly expanding program in Anchorage, the largest population center in state. Research opportunities, summer teaching, plus chance to help formulate departmental programs and policies. Write to: William E. Bedsworth, Coordinator, M.B.A. Program, University of Alaska, Southcentral Regional Center, 1820 West Northern Lights Boulevard, Anchorage, Alaska, 99503.

Economics: September, 1970; Ph.D. required. Teaching assignment flexible, including range from elementary to advanced undergraduate courses and/or graduate courses. Rank and salary commensurate with qualifications. T.I.A.A., life insurance, major medical, other fringe benefits. Please send résumé to: Dr. Charles L. Kraemer, Chairman, Department of Economics and Business Administration, Wagner College, Staten Island, New York, 10301.

Economics: The United Nations has a continuing need for economists for challenging research and operational posts at Headquarters, New York, in Geneva, Vienna, and in developing countries throughout the world. These posts cover the full field of economic specialization (econometrics, fiscal policy, foreign trade, finance, agriculture, mining, water resources, industrial management, development, urban planning, etc.), at all levels of responsibility and of varying duration—permanent or short term. Representative of the posts now under recruitment are the following. The gross salary for these three positions ranges between \$12,380-\$20,490 (\$9,666-\$15,118 tax free), exclusive of dependence and post allowances which may amount to an additional several thousand dollars. (1) UNCTAD, Geneva: econometrician experienced in econometrics of primary commodity markets and/or international trade. (2) UNIDO, Vienna: economist experienced in industrial development strategies, plans and projections as well as industrial project preparation and implementation. (3) Headquarters, New York: Minerals economist, preferably with a doctorate and five years of practical experience in the field; fluency in French essential. U.S. citizens interested in obtaining additional information about these as well as other U.N. employment possibilities are requested to send a résumé to: Office of International Organization Recruitment, Room 4336A (IO/IR), U.S. Department of State, Washington, D.C., 20520.

Industry model building: The development of creative approaches to forecasting industrial performance. A strong background in econometrics is required. The position is with the Investment Technology Group of the Boston Company doing

state-of-the-art research in scientific investment decision making. Send résumé, including salary requirements, to: Richard C. Katz, Boston Company, 100 Franklin Street, Boston, Massachusetts, 02106. Telephone 617-542-9450.

Business administration: Rapidly expanding business program of a medium-size state college located in a residential community, offering B.A. in Economics and B.S. in Business Administration in the fields of accounting, economics and management, has a position open at the professor rank. Courses to be taught include principles and undergraduate electives (negotiable) with possibility of teaching graduate level courses. Requirements: D.B.A. or Ph.D. in business administration and college teaching or allied experience. Salary is commensurate with qualifications. For further details, write to: Dr. Joseph M. Thorson, Chairman, Department of Economics and Business, West Chester State College, West Chester, Pennsylvania, 19380. Telephone 215-436-2236.

Accounting: Rapidly expanding business program of a medium-size state college located in a residential community, offering B.A. in Economics and B.S. in Business Administration in the fields of accounting, economics and management, has a position open at assistant or associate professor rank. Courses to be taught include principles and undergraduate electives (negotiable) with possibility of teaching graduate level courses. Requirements: C.P.A. or master's degree in accounting, and college teaching or other allied experience. Salary is commensurate with qualifications. For further details, write to: Dr. Joseph M. Thorson, Chairman, Department of Economics and Business, West Chester State College, West Chester, Pennsylvania, 19380. Telephone 215-436-2236.

Marketing and/or retailing: Rapidly expanding business program of a medium-size state college located in a residential community, offering B.A. in Economics and B.S. in Business Administration in the fields of accounting, economics and management, has a position open at the assistant or associate rank. Courses to be taught include principles and undergraduate electives (negotiable) with possibility of teaching graduate level courses. Requirements: Ph.D. required; would consider a candidate at dissertation stage. Salary is commensurate with qualifications. For further details, write to: Dr. Joseph M. Thorson, Chairman, Department of Economics and Business, West Chester State College, West Chester, Pennsylvania, 19380. Telephone 215-436-2236.

Economics and/or finance: Rapidly expanding business program of a medium-size state college located in a residential community, offering B.A. in Economics and B.S. in Business Administration in the fields of accounting, economics and management, has a position open at the assistant or associate rank. Courses to be taught include principles and undergraduate electives (negotiable) with possibility of teaching graduate level

courses. Requirements: Ph.D. required; would consider a candidate at dissertation stage. Salary is commensurate with qualifications. For further details, write to: Dr. Joseph M. Thorson, Chairman, Department of Economics and Business, West Chester State College, West Chester, Pennsylvania, 19380. Telephone 215-436-2236.

Economist: Senior appointment in macro- or microeconomic theory. Person with substantial publication record and doctoral dissertation experience who would play active role in department's expansion from master's to doctoral program. Opening for September, 1970. Salary open and competitive. State university located in Midwest. P364

Economist: Full professor, associate professor, assistant professor or instructor. Salary \$7,500-\$13,400. Ph.D. plus teaching experience preferred but will consider D.B.A. Beginning September, 1970. Prefer person with background in general theory, money and banking, and some quantitative ability. No publishing necessary; strong commitment to teaching mandatory. Send résumé to: Dr. John Creaghe, Dean, Social Science Division, New England College, Henniker, New Hampshire, 03242.

Accounting: Ferris State College is looking for an accounting professor; rank and salary dependent upon experience and background of candidate. Salary can go to \$17,000-plus with the proper background. This can be a very attractive position for a person with excellent credentials. Write or phone: Dr. Lowell Chapman, Head, Accounting Department, Ferris State College, Big Rapids, Michigan, 49307.

Management, finance, insurance, statistics: Ferris State College has positions available in management, finance, insurance, and statistics for 1970-71. Two positions in management with emphasis upon organization theory or policy or both. The finance and insurance positions involve new programs. Appointments will be made at various levels, depending upon education and qualifications. Ph.D. or D.B.A. plus experience desirable. Will consider applicants near the degree. Stress is on teaching. A new Bureau of Business Services will begin operating in June, 1970. The School of Business will move into a new building in March, 1970. The appointee will have an opportunity of engaging in the activities of the Bureau. Address applications to: Dr. Alex S. Pomnichowski, Chairman, Department of Business Administration, Ferris State College, Big Rapids, Michigan, 49307.

Business administration: Small Gulf Coast four-year liberal arts college has vacancy for man to offer courses in marketing, corporation finance, management, production, investments. Beginning September 1, 1970. Doctorate preferred. Salary negotiable, depending on qualifications. Write to:

Samuel M. Betty, Acting Chairman, Commerce Department, Spring Hill College, Mobile, Alabama, 36608.

Management: Ph.D. in economics or D.B.A. (or candidates who are close to attaining one or the other of these degrees) sought for fall, 1970, to teach in three-year-old, rapidly growing Master of Business Administration Program, as well as at the undergraduate level in business administration. Salary and fringe benefits fully competitive with other institutions; rank depends upon experience. For further information, make application to: W. P. Carton, Director, Master of Business Administration Program, Loyola College, 4501 North Charles Street, Baltimore, Maryland, 21210. Telephone: 301—435-2500.

Marketing: Ph.D. in economics or D.B.A. (or candidates who are close to attaining one or the other of these degrees) sought for fall, 1970, to teach in three-year-old, rapidly growing Master of Business Administration Program, as well as at the undergraduate level in business administration. Salary and fringe benefits fully competitive with other institutions; rank depends upon experience. For further information, make application to: W. P. Carton, Director, Master of Business Administration Program, Loyola College, 4501 North Charles Street, Baltimore, Maryland, 21210. Telephone: 301—435-2500.

Statistics, quantitative methods in business: Ph.D. in economics or D.B.A. (or candidates who are close to attaining one or the other of these degrees) sought for fall, 1970, to teach in three-year-old, rapidly growing Master of Business Administration Program. Salary and fringe benefits fully competitive with other institutions; rank depends upon experience. For further information, make application to: W. P. Carton, Director, Master of Business Administration Program, Loyola College, 4501 North Charles Street, Baltimore, Maryland, 21210. Telephone: 301—435-2500.

Economist: Position as assistant professor in expanding economics department beginning fall, 1970. Ph.D. required; fields other than international, development, thought and public finance. Salary \$11,000 and up for academic year. T.I.A.A. Located in vacation/sports area of upper Midwest. Please send résumé to: Chairman, Department of Economics, Northern Michigan University, Marquette, Michigan, 49855.

Economists: The Indian Institute of Technology/Kanpur, Kanpur, U.P., India, has openings for all ranks from lecturer to professor. Preference given to candidates capable of doing analytical work; proficiency in use of quantitative techniques, having strong background in mathematics-statistics; Ph.D., strong academic career; ability to do creative research. IIT/K is assisted by the U.S. government and ten leading U.S. educational institutions. Benefits include transportation to

India. Airmail applications stating academic references to: Professor K. K. Singh, Chairman, Department of Humanities and Social Sciences, Indian Institute of Technology/Kanpur, IIT Post Office, Kanpur, U.P., India.

Economic theory, statistics, finance: Openings for full and associate professors. Requirements are Ph.D., D.B.A., or M.B.A. with course work completed for doctorate and college teaching, business, or government experience. Publications desirable but not essential. Salary range depending on education and other qualifications; nine months with summer session optional; annual increments; high fringe benefits; excellent possibilities for advancement. A Catholic Liberal Arts College with a Division of Business, located in pleasant suburban surroundings in up-state New York. Send résumé to: Chairman, Business Division, Siena College, Loudonville, New York, 12211.

Economist: Department chairman for fall, 1970. Private midwestern college has opening for a man with Ph.D. in economics. Rank and salary open according to education and experience. Should be capable of directing a ten-man department plus graduate teaching. Please submit complete résumé to: Dean William J. Hoben, School of Business Administration, University of Dayton, Dayton, Ohio, 45409.

Economists: Several permanent positions in private midwestern university, starting fall, 1970. Rank and salary open according to education and experience; teaching experience, research experience, and publications regarded as especially valuable. Please submit complete résumé to: Dean William J. Hoben, School of Business Administration, University of Dayton, Dayton, Ohio, 45409.

Economists: New York State has openings in various departments for general economists and economists with specialization in labor, business, taxation, and transportation economics. Starting salaries \$12,585 or \$15,590, depending on education and experience. Write: N.Y.S. Department of Civil Service, R-876A, Albany, New York, 12226.

Chairman, department of business administration: Morehead State University is seeking a department chairman for business administration, July 1 or September 1, 1970. The Department of Business Administration has the following undergraduate programs: accounting, finance, marketing, management, and data processing. An M.B.A. program is presently under study. The appointee will have a small teaching load in addition to his duties as department chairman. A doctorate in business administration (Ph.D. or D.B.A.) is preferred but a doctorate in education will be considered. The salary range is between \$16,500 to \$17,500, depending on experience and qualifications. Send a résumé to: Mr. Alex D. Conyers, Chairman of the Screening Committee, School of Applied Sciences and Technology, Morehead State University, Morehead, Kentucky, 40351.

Computer sciences, statistics: Ph.D. or near-Ph.D. desired. Doctoral preliminaries minimum requirement. Rank and salary open for negotiation. Send personal data sheet to: Professor James E. Jensen, Chairman, Department of Economics and Business, Wisconsin State University, Stevens Point, Wisconsin, 54481.

Structure of industry, business and government: Ph.D. or near-Ph.D. desired. Doctoral preliminaries minimum requirement. Rank and salary open for negotiation. Send personal data sheet to: Professor James E. Jensen, Chairman, Department of Economics and Business, Wisconsin State University, Stevens Point, Wisconsin, 54481.

Marketing or accounting: Ph.D. or near-Ph.D. preferred; M.S. minimum. Rank, salary open for negotiation. Send personal data sheet to: Professor James E. Jensen, Chairman, Department of Economics and Business, Wisconsin State University, Stevens Point, Wisconsin, 54481.

Quantitative economist (management science): Growing, high-quality graduate program in industrial administration and operations research has opening for young economist with good training in mathematics, management science, and computer modelling. Ph.D. required. Send résumé to: Alfred L. Thimm, Director, Industrial Administration, Industrial Administration Institute, Union College, Schenectady, New York, 12308.

Mathematical economics, econometrics, microeconomic theory: Three positions are available. Applicants must have Ph.D. by September, 1970. Appointments will be made at the rank of assistant or associate professor within a salary range of \$10,500 to \$13,700, depending upon teaching experience and qualifications. Send application to: William C. Hoekendorf, Chairman, Department of Economics, Eastern Washington State College, Cheney, Washington, 99004.

Administration: Dean of School of Business and Economics. Evidence of scholarly achievement

such as a doctorate is required in business or economics, preferably with background in both fields and administrative experience in the combination of fields. Should be eligible for rank of professor. Salary range from \$19,404 to \$23,604. Beginning July 1-September 1, 1970. Inquire: Milton Dobkin, Vice President for Academic Affairs, Humboldt State College, Arcata, California, 95521.

Economists: Seeking either a single economist or group of economists who are experienced in systems analysis in order to build a computerized program of trading commodities. The salary is high and there are additional bonus considerations. Kindly submit your résumé to: Marvin Friedman, Room 736, 60 East 42nd Street, New York, New York, 10017.

Dean: Appalachian State University announces plans to establish a College of Business on July 1, 1970. The University is a part of the system of public higher education of the state of North Carolina. It offers the B.A., B.B.A., M.A., and M.S. degrees. It has 6,833 students. Of these, 911 are undergraduate majors in business and 48 are graduate majors in business. At the present time a search is being conducted for a person to serve as dean of this College. Nominations or applications should be addressed to: Dr. Paul Sanders, Provost, Appalachian State University, Boone, North Carolina, 28607.

Economists: Many openings during the year, calling for all specialties and all levels, instructor to department chairman. Senior positions for faculty with high professional standing and long experience. Candidates provided broad visibility, yet confidentiality protected through referral only to positions which match candidate's qualifications. Salaries competitive. Ph.D. with teaching experience preferred. Free service provided by nonprofit organization serving over 300 liberal arts colleges and universities throughout the U.S. Write for one-page registration form and further information to: Cooperative College Registry, 1 Dupont Circle, N.W., Washington, D.C., 20036.

Economists Available for Positions

Italics indicate fields of specialization

Economic theory, monetary and fiscal theory and policy, international economics, mathematics for economists, history of economic thought, comparative economic systems: Man, 27; A.B., A.M. from major universities with total of 60-plus semester units of graduate study as of September. Two years of experience teaching principles. Desires position in 4-year liberal arts college within commuting distance of major university for further work toward Ph.D. in economics; further study will include quantitative methods, statistics, econometrics, advanced mathematics. E1967

International development, economic history, microeconomic theory, industry and government,

economic policy: Man, 26, married; B.A., M.A., Ph.D. candidate at leading New England university. Background includes 14 different college courses taught, experience in course development, departmental organization and faculty-student-administration relations. Presently department chairman at small liberal arts college. Prefers teaching position with graduate program in New England or Middle Atlantic States area.

E1988

Urban economics, U.S. and foreign economic development, housing market: Man; A.B., M.S. Thirty-eight years of high-level practical experience; widely traveled; knowledge of Spanish and

Italian. Current security clearance. Fully-equipped private office in Washington, D.C. Available for a limited number of additional assignments on a contract, retainer, or per diem basis. Brochure on request. E1994

Micro, development, international principles, history of economic thought: Man; B.Sc. (Econ.), Hons., London, Experienced teacher and research worker. Author; free lancer. Recent work at research institute, London. Available immediately. Syed Ahmad Ali, Guzri, Patna 8, India.

Economic thought, labor problems and law, collective bargaining, principles: Man; Ph.D., eastern university. Extensive teaching experience; excellent references; book in process. Seeks position in Middle Atlantic States area. E1995

Principles of economics, micro, welfare, history of economic thought: Man, 30; Ph.D. Five years of university teaching. Seeks teaching position with opportunities for research in southeastern U.S. Available in June or September, 1970.

E1998

Micro- and macroeconomic theory, statistics, econometrics, business conditions analysis, industrial organization: Man, 25, married; B.A., Ph.D. candidate with all course work and preliminary examinations completed. Presently devoting full time to dissertation under Woodrow Wilson Dissertation Fellowship; expect completion by summer, 1970. Desires teaching position in Midwest. Available in September, 1970. E2008

Economic development, Latin-American economic problems, agricultural development, micro- and macroeconomic theory, marketing, principles: Man, married; Ph.D., Harvard. Extensive experience teaching, research, and consultant. Numerous publications, including book. Now professor at small eastern college. Wishes to relocate in position involving teaching, research, and/or consulting. Available in June or September, 1970. E2011

International economics and administration, planning and project evaluation, finance, Latin-American development: Man, 39, married; Ph.D. Government, business, and university research and teaching experience in U.S. and Latin America, including grants. Currently in second year as Fulbright-Hays Professor in Mexico, research and teaching. Seeks position with tenure in university teaching, research, and possibly administration. Available after June, 1970. E2012

Price theory, mathematical economics, macroeconomics, computer application in economics, market structure, agricultural economics, international economics, economic development: Man, 35, married, Indian national; M.S., Ph.D., U.S. state university, recent Ph.D in agricultural economics. Three years of college teaching experience; three years of research in university. Current

recipient of a national foundation grant; modest publications. At present an assistant professor of economics in a U.S. state college; wishes change for location improvement. Desires teaching and research at college or university. Prefers in or near Washington, D.C.-Baltimore metropolitan area but will consider other offers. Will consider government position or position in international organizations. Résumé furnished upon request. Available in June, 1970. E2013

Economic theory, international economics, growth and development, price theory, agricultural policy, African economics: Man, 31, married; B.A., Cornell, M.B.A., Wharton, Ph.D., Oxford, England (economics). Four years of teaching experience; five years of research experience, including two in Central Africa; in international control of trade, subsistence agriculture and national marketing boards. Desires research and/or teaching position located in North or West Coast of U.S. or Canada. Available in August, 1970. E2021

Principles of economics. Man, 44, married; M.A., all course work completed for Ph.D. Eight years of college teaching experience, both graduate and undergraduate classes; several years of business experience. Publication and research experience. Special interest in heading a principles program. Available in January, 1970. E2031

Micro- and macroeconomics, money and banking, international economics: Man, 47; B.A., M.A., Ph.D. Experience includes five years of teaching and chairman of department; directing and conducting research; Federal Reserve Board staff; and management consultant to banks. Publications. Current position very satisfactory (tenure and full professor) but wishes to relocate. Desires teaching position. Would consider chairmanship. Available in September, 1970. E2033

Principles, theory, history of economic thought, money and banking, American economic history: Man, 42, married; Ph.D. Fifteen years of teaching experience; administrative experience, including department chairmanship; some publications. Desires position in business, teaching, or research. E2034

Natural resource economics, price analysis, cost-benefit analysis, public finance, public policy, agricultural economics, quantitative methodology: Man, early 40's, married; Ph.D. Six years of experience at assistant and associate professor levels, currently head of a price analysis division in the federal government. Desires teaching, research, and/or extension position including administration in a non-urban area of northern United States or Canada. Available on reasonable notice. E2036

Micro- and macroeconomics, demography: Man, 33, married, Pakistani; B.A., M.A. Six years of research experience, mostly in the field of demog-

raphy. Modest publications and list of work assisted. Excellent references; an intended immigrant to the U.S. (immigrant visa approval has already been obtained). Desires assistant's position in teaching/research either on temporary or permanent status in any kind of organization. Available on two months' notice. Résumé on request. E2038

Management science, capital budgeting, accounting microeconomics: Man, 42, married; B.S. (engineering), C.P.A., M.S., Ph.D. (business and economics) expected dissertation completion 1970. Eleven years of part-time and some full-time teaching of quantitative economic and management science courses with twenty-two years of industrial and business experience. Recent five years as head of operations research. Desires teaching and/or administrative position in industrial management science function. Prefers mid-west or western location. E2043

Theory and practical statistics, transport, international trade, public finance: Woman, 29; Ph.D., Indian national. Head of Economics Department for a degree college; five years of teaching experience. Desires a position with a progressive institution. Available for an interview. E2046

Cooperative marketing of agricultural produce, international trade, international aid, rural development: Man, 30; Ph.D., Indian national. Field officer for the United Nations development program in India. Formulated development projects for economic and social development of rural and urban areas. Seeks position anywhere in the U.S. or anywhere in the world. E2047

Econometrics, mathematical economics, economic theory, economic development and growth, regional and international economics: Man, 36, married; M.S., Ph.D. (economics, statistics). Experience in graduate and undergraduate teaching. Publications. Desires relocation to a teaching and/or research position. E2050

Labor economics, trade unionism, industrial relations, collective bargaining: Man, 35; M.A., Ph.D. Nine and a half years of university teaching experience. Research experience and published research papers. Desires any position in teaching or research on fulltime basis; contract job. E2051

Economic problems of underdeveloped areas, international economics, labor economics, microeconomics, input-output analysis: Man, 42, single; doctor of economics, Geneva University. Seven years of experience in teaching and economic research in Europe and Canada; publications in English and French. Seeks teaching or research position in U.S. E2054

Public economic policy, introductory economics, finance: Man, 61; Ph.D. (economics, political science, law), Wisconsin. Retiring from thirty-

two years of federal executive and legislative service, largely in policy research; economist in top grades since 1945. Early favorable experience in undergraduate and graduate teaching (Wisconsin, Texas, Cincinnati) reinforced by staff nature and public communication orientation of principal federal responsibilities. Interested in limited teaching schedule, seminar preferably included, or public policy research in Washington area; also in single-term faculty appointments and campus-based Indian reservation development programs elsewhere in U.S. Available in January, 1971. Roderick H. Riley, 7021 Richard Drive, Bethesda, Maryland, 20034.

Accounting, finance, management: Man, 36, married; M.B.A., C.P.A., Ph.D. Experiences include seven years of teaching accounting and management, eight years of public, industrial, and government accounting. Some publications and EDP background. Interested in research and publications with teaching. Available in September, 1970. E2056

Economics of developing regions: Man, 46, married; B.S., advanced degree dissertation in developing country economics in progress. Papers on population growth in developing country and on international finance subjects published. Twenty-two years of extensive experience in senior finance position in national transport, eight of which as officer in charge. Five years as management auditor and project activity evaluator on financial and economic analyses of national development projects financed by U.S. government agency in fourteen countries. Proven diagnostic ability. Excellent references. Unusual experience overseas. Seeks position offering creative challenge. E2058

Macro- and microeconomic theory, economic development, international trade, statistics: Man, married; Ph.D. Six years of teaching experience in dissertation supervision; publications. Available in September, 1970. E2061

Economic history, microeconomics, history of economic thought, labor economics: Man, 26, single; A.B., M.A., Ph.D. dissertation in progress to be completed by fall, 1970. Four years of teaching experience; fellowships; some publications. Desires teaching position in U.S. or Canada. Available in fall, 1970. E2063

Principles, money and banking, theory: Man, 33, married; B.A., M.A., 33 graduate units towards Ph.D. Woodrow Wilson Fellow. Two years as teaching assistant; two years of lecturing at a four-year liberal arts college in Canada; four and one-half years in economic research and practical banking with a large commercial bank. Desires to return to teaching. Seeking position at a junior college or liberal arts college within commuting distance of a university with a Ph.D. program in economics. E2066

Macro- and microeconomics, regional economics, natural resource and land economics: Man, 34, married; B.A. (English), Master of Forestry, Ph.D. dissertation in progress. One year government research; two years of part-time teaching. Desires teaching position. Résumé furnished on request. Available in June, 1970. E2067

Urban and regional economics, statistics, econometrics, human resources, medical economics, public finance, micro- and macroeconomic theory: Man, 30, married, U.S. permanent resident; M.A., Ph.D. dissertation in progress and degree expected soon. Variety of awards and fellowships. Four years of quantitative and industrial research experience in foreign countries and the U.S., in government and nonprofit organizations. Teaching experiences in statistics and economics. Publications, including academic journal articles; input-output, economic base studies for a major U.S. standard metropolitan area; demographic and transportation projections. Desires a university teaching position with research opportunity or a research position in nonprofit organization, private firm, or government in the above-mentioned areas requiring quantitative background. Will relocate. Résumé furnished upon request. Available immediately. E2068

International trade and finance, economic theory and thought, economic development, Western and Soviet-type economic systems, money, labor: Man, 54; Ph.D. Ford Foundation research fellowship; publications (one book). Now on a state university faculty. Desires a teaching position with a progressive institution. Available in September, 1970. E2069

Accounting, business administration, business law: Man, 55; C.P.A., lawyer, LL.M., Ph.D. Desires teaching-administrative position. Substantial teaching, government, and business experience. Résumé on request. E2071

Macro- and microtheory, international trade and development, quantitative methods: Man, 28, married, European (permanent resident); B.A., M.A., Ph.D. candidate, dissertation in progress. Speaks English, French, Italian, and some Spanish and Arabic. Seeking administrative, research and/or consulting position. Résumé available upon request. Available in July, 1970. E2072

Mathematics for economists, microeconomics, macroeconomics, comparative economic systems: Man, 35; Ph.D. More than six years of university teaching experience; publications. Seeks teaching post or research fellowship. Available in September, 1970. E2074

Macro- and microtheory, regional development, operations research, simulation models, philosophy of science: Man, 42; Ph.D., Stanford. Fifteen years of teaching, research, and administrative experience. Excellent teaching record; publications. Four years as study director of large-

scale research projects. Wishes to return to academics. Seeks senior teaching/research position or deanship/chairmanship with some teaching. Excellent rapport with students and faculty. Prefers western states. E2077

Finance, management science, planning: Man, married; undergraduate degree in chemistry and mathematics; Ph.D. (business administration). Extensive diversified teaching; business experience as business manager, purchasing agent, and consultant; currently professor and chairman of business division. Desires teaching and/or administrative position at undergraduate or graduate level; also interested in corporate planning and management development. E2080

Economic development, economic history, money and banking, international economics, history of thought: Woman, 39, separated; B.S., M.A., all requirements for Ph.D. completed except dissertation. Experience in university teaching and research. Fulbright and other awards. No strong geographical preference. Available for teaching and research in fall, 1970. E2084

Money and banking, international economics, labor problems, economic problems of Latin America, economic principles: Man, U.S. citizen; Ph.D. Ten years of teaching experience in economics on college level. Will relocate in the U.S. or overseas. Available in summer or September, 1970. E2085

Economic problems of developing countries, population problems, economic development and planning, fiscal policy: Man, 25, single; B.Sc., B.A., M.A. Good references; an intended immigrant to the U.S. (already received approved Department of Labor certification). Desires assistant's or any other position in teaching or research either on temporary or permanent basis in any school, college, research institution, or any other kind of organization anywhere in the U.S. with or without opportunities toward Ph.D. (Migration possible only with a job or research offer.) Available at a month's notice. E2092

Economic theory, statistics and decision theory, forecasting, finance: Man, 46; Ph.D. Nine years of teaching experience plus business experience; present position associate professor at large midwestern university. Available in summer or fall, 1970. E2093

Manpower management, industrial relations, personnel administration, organization theory: Man, 28; two master's degrees (manpower planning and public administration), Ph.D. (industrial relations), University of Minnesota. Seven years of research and teaching experience. Seeking research and/or teaching position. Available in September, 1970. E2095

Manpower economics, collective bargaining, regional-urban development economics, micro- and

macrotheory, money and banking, principles: Man, 28, single; B.A., M.S., plus twenty semester hours, all from recognized schools, presently enrolled in Ph.D. program. Liberal arts orientation. Four years of full-time college and university teaching experience. Some administrative experience. Desires teaching position or position in government agency. Résumé upon request. Available in September, 1970. E2096

Macro- and microeconomics, public policy, philosophy of economics, managerial economics, economic history, general economics: Man, 30, married; B.S., M.S. plus forty-five. Four years of college teaching; National Science Foundation fellow; currently assistant professor in large four-year college in New York State; former chairman of business administration and economics; interested in natural resource economics and public policy. Seeks teaching position in two- or four-year college in Denver or Pacific Northwest. Desires change for location improvement. Available in June or September, 1970. E2097

International trade and finance, economics of developing countries, economic planning, micro- and macroeconomics, economic systems, principles of economics: Man, 40, naturalized Canadian; Ph.D. Ten years of university teaching and consulting experience; currently chairman of a ten-member department. Some publications. Desires relocation. E2098

Urban and regional economics, economic growth and development, air transportation: Man, 34, married; M.B.A. in economics (statistics minor), University of Chicago, with academic honors. Ten years of professional experience in economic planning and research projects with travel organization (economic analyst), regional planning commission (research associate), and consulting firms (senior economist and director of economic research). Desires overseas position in Far East, Europe, or East Africa with international organization or private firm. Numerous project reports available on request. E2102

Public and corporation finance, international trade, business law, taxation, national accounting: Man, 29, married, Portuguese; graduate courses in economics, from the University of Lisbon. Research fellow and assistant since 1964 at the same university. Substantial publications. Seeks American organizations interested in consulting and/or representation in Portugal. E2103

Public control, economic development, corporate and long-range planning, macroeconomics, business organization and management: Man, 60; B.S., M.S., Ph.D. Teaching experience at two Big Ten universities. Career as an economist; federal government policy official; corporate planning executive. Available in early 1971 for administration or teaching in college or university. E2104

Economic theory, money and banking, government and business, management, finance: Man, 57; A.B., A.M., University of Pennsylvania, Ph.D., New York University. Extensive college teaching and administrative experience as department chairman and director of a school of business administration. Some modest publications. Presently teaching in Midwest but would like to relocate in South. Prefers chairmanship of department where teaching duties expected but am interested in any challenging building job. Complete résumé on request. E2105

International economics, economic development, economic theory: Man, 48; M.B.A., M.A., Ph.D., University of Chicago. Phi Beta Kappa; fellow, London School of Economics, 1962-64. Considerable experience in academic, government, and research activities. Numerous publications, including book, journal articles, review, reports, etc. Professionally abreast of developments, including mathematical and statistical techniques. Active research interests. Nonprofessional interests include music, art, other activities. Visiting scholar, Brookings Institution, 1967. Fluent in Spanish. Seeks high-level, intellectually progressive placement. E2106

International economics, micro-macro theory, economic development, comparative systems: Man, 29, married; Ph.D., University of California, Berkeley. Teaching and research since 1967. Five recent publications in international trade and development. Research interest in input-output techniques and regional economics. Desires teaching and/or research position. Available in September, 1970. E2107

General resource economics, economics of outdoor recreation, general forest economics: Man, 29, U.S. citizen; Ph.D., 1968. Publications. Presently employed in natural resource economics research for a foreign government. Seeks position in teaching and/or research. E2108

Comparative systems, public finance, international economics, economic development: Man, married; B.S., M.A., Ph.D. Five years of college teaching; one year in private industry; consulting experience. Publications. Currently at southern state university. Desires teaching position with opportunity for research. Available in June or September, 1970. E2109

Economic development, microeconomics, economies of Africa, statistics: Man, 36, married; B.A., Ph.D. from leading university. Ten years of research and teaching at graduate and undergraduate levels. Currently associate professor of economics at large eastern university. Desires to leave academic life and enter antipoverty work with social service agency, government (preferable local or state), or corporation with a social action program. Not interested in analysis unrelated to practical impact of program. E2110

International economics, agricultural development, industry: Man, 37, married; B.A., U.C.L.A., plus year and a half graduate work. Has several languages. Eight years of diversified experience in U.S. and overseas, including positions involving public relations; now supervising research on industrial location and evaluation. Résumé furnished upon request. Desires position with private firm; will relocate. E2111

International economics, macro- and microeconomics, money and banking, public finance, principles of economics, labor economics, business organization: Man, 27, married; B.S., M.A., University of Nebraska, two years of additional graduate work, University of New Mexico. One year of teaching experience. Publications in the field of state and local finance. Desires teaching in liberal arts college. Available in July, 1970. E2112

Economics, marketing: Man, 33, married; B.A., B.S., M.A., currently working on Ph.D. dissertation. Six years of teaching experience; three years of research experience in natural resource economics and marketing. Several publications. Interested in research, consulting, or administrative position with university, government, or private firm. Prefers Middle Atlantic and Southeastern States. E2113

Latin American affairs, urban and regional development planning, regional analysis, political and institutional sociology of development: Man, 42, married; Ph.D. expected in June, 1970. Four years as adviser in regional development in Latin America. Teaching experience. Publications. Desires teaching and/or research position. Prefers Rocky Mountain, East or West Coast U.S. or Canada. E2114

Development planning, transport programming, project analysis and evaluation, development finance: Man, 56; U.S. citizen; only thesis incomplete for Ph.D. Over twenty years of experience in international organizations and private consulting, involving advisory, operational, and teaching experience in above fields in Far East, Latin America, and Eastern Africa. Fluent Spanish, good Portuguese and German, some French. Seeks economic consulting, possibly combined with part-time teaching, on problems of economic development planning, project preparation and evaluation, and financial implementation in less developed countries. Prefers location in Washington or San Francisco areas; in Mexico, Central America, or Puerto Rico; or in Western Europe. E2115

Public finance, micro- and macroeconomics: Man, 34, married; Ph.D. Six years of teaching and research experience. Some publications. Desires teaching position with opportunity and facilities for research in the area of public finance, particularly state and local finance. Available in July, 1970. E2116

Comparative economic systems, economics of the Soviet Union, micro- and macroeconomics, labor economics, history of economic thought: Man, 44, married; Ph.D., with honors. Fourteen years of teaching experience; presently associate professor and chairman of economics department. Seeks academic position with chance to combine teaching and research. Available in September, 1970. E2117

Development, international economics, regional economics, micro and macro: Man; U.S. citizen; Ph.D. Currently associate professor and department chairman in a middle eastern university. Available in September, 1970, or January, 1971. E2118

Micro, macro, history of economic thought, economic history, monetary theory, public finance, economic development, agrarian economics: Woman, 38; Indian national; Ph.D., London School of Economics. Thirteen years of teaching and research experience in India, London, U.S., and Canada. Currently with the government of India. Modest publications in American journals. Available for permanent faculty appointment in U.S. in 1970. E2119

Microeconomic theory, econometrics, statistics, regional economics, mathematical economics: Man, 31, married; B.S. (mathematics), Ph.D. (economics). Three years of teaching and research experience. Desires teaching and research position in western U.S. or Canada. Available immediately. E2120

Economic theory, econometrics, systems analysis, industrial organization, welfare economics: Man, 28, married; B.A., M.A., M.S. (M.I.T.). Fluent French; publications in health and medical care fields. Three years of systems analysis and programming master corporate production planning model, totally on-line; market and industry analyses; input-output analysis for U.S. economy with private research organization. Research and work experience with federal, state, and municipal agencies in economic planning. Desires work in economic planning and resource management. Available in September, 1970. E2121

Public finance, money and banking, labor, theory: Man, 49, married; Ph.D. Five years of university teaching experience, committee work, and program development. Ten years of complex administrative and management experience. Desires chairmanship of growing and progressive economics department. E2122

Economic principles, international and economic development: Woman, 28; B.A. (honors), M.A. Four and a half years of college teaching experience. Desires teaching position. Available in September, 1970. E2123

Finance, banking, macro- and microeconomics, international finance and business, marketing,

accounting: Man, 38; B.A. from England, M.S., Columbia University, Ph.D. expected in 1970. Nine years of teaching and business in the U.S. and three abroad. Available in September, 1970. E2124

Statistics, international economics, money and banking, programming: Man, 27, single; M.A., Ph.D. dissertation in progress. Currently in third year of teaching at large university; one and a half years of programming experience. Résumé available. Desires position with private firm; will relocate. Available in June, 1970. E2125

Securities, money and banking, economic development, accounting: Man, 31; M.A. in Economics. Broad background which includes teaching, international trade, trust company and research experience. Extensive overseas exposure; fluent in four languages. Desires teaching or economic research position. Will consider assignment abroad or traveling. Available in June or September, 1970. S. Totah, Box 1374, Station B, Montreal 110, Canada.

Monetary theory and policy, statistics and econometrics, urban economics: Man, 28, married; Ph.D. Six years of undergraduate teaching and research experience. Prefers university and/or research position in the Northeast. E2126

Cost analysis economics of defense or transportation: Man; M.B.A., Harvard. About twelve years of economics and transport analysis in Europe, Asia and another twelve years of weapons systems and transport analysis in U.S. Equally interested in U.S. or overseas location. Desires to continue managerial or project leader responsibility utilizing systems analysis experience. E2127

Urban-regional economics, money and banking, international economics, comparative systems: Man, 26, married; A.B., M.S. Ford Foundation Grant in Public Affairs. Two years of teaching at college level. Currently consulting economist to federal government. Seeks research or business position in West (preferably California). E2128

Financial institutions, cooperatives, business administration: Man, 35, married; in the process of getting Canadian citizenship; law degree. M.S., Columbia University. Broad background; eleven years of diversified experience: foreign service (commercial secretary), South America; central bank, university, investment fund. Currently head of the financial statistics division of a Canadian provincial government. Fluent in French, Spanish, English; some knowledge of Italian. Seeks research or administrative position outside of North America (either in Latin America or in Europe) with an international organization or private industrial or financial firm. Present salary \$14,500. Available on reasonable notice. E2129

Labor economics, labor history, micro- and macroeconomic theory, principles of economics, current economic policy: Man, 31; all requirements for Ph.D. have been completed with the exception of the dissertation, which is being revised now. Seven years of college teaching experience. Desires position at academic institution having economics as a major. Preferred location would be east of Indiana. Available in September, 1970. E2130

Business finance, investments: Man, 50; Ph.D. Fifteen years of teaching experience. Interested in a teaching position in the Midwest or Southwest. Presently employed but available in June or September, 1970. E2131

Economic theory, public finance, urban economics, monetary theory, international economics, history of economic thought, accounting: Man, 26; B.B.A., M.A., Ph.D. candidate, completing comprehensives and preparing to begin dissertation. Scholarships and other academic recognitions. Background includes five years of teaching experience. Seeks position on the college level as of September, 1970. E2132

Agricultural economics, economic theory, econometrics: Man, 25, married; B.S. University of Wisconsin, in final stages of preparing Ph.D. dissertation at University of California, Berkeley, in agricultural economics in area of production economics and econometrics. Interest also in problems of public policy, rural poverty, rural-urban adjustment, and problems of environment. Desires research-teaching position in university although will consider other types of positions. Available in summer, 1970. E2133

Manpower, labor economics, industrial relations, labor history, economic history: Man, 57, married; Ph.D., eastern university. Eight years of teaching experience; administration and research experience; numerous publications. Now research consultant to U.S. Congress on manpower and labor problems. Desires teaching position with research opportunities in manpower and labor field in South or Midwest. Available in September, 1970. E2134

Federal income taxation and public finance, public utility economics, corporation finance: Man, in 50's; Ph.D. Background combines academic training with extensive experience with IRS and Treasury and regulatory agencies. Teaching, administrative, or consulting position desired in California area. E2135

Micro- and macroeconomics, money and banking, welfare, principles, labor economics, public finance: Man, 49; M.A. (philosophy), Ph.D. (economics). Fifteen years of college teaching; eight years as department head. Prefers liberal

arts college and interdisciplinary opportunities. Currently serving sentence in state prison system for burning of draft files. Available in September, 1970. E2136

Economic theory: Man, 25, single; B.A., M.S., plans to begin Ph.D. course work in two years. Has had three years of experience of applied microeconomic theory in the planning departments of two large midwest corporations; supply analyst; excellent public speaking ability (former collegiate debater) and well-developed writing skills. Prefers West Coast. Available in June or September, 1970. E2137

Economics of transportation, business, transportation and maritime law, traffic management, law of freight loss and damage claims, I.C.C. rules, export and import procedures: Man; Ph.B., M.A., LL.B. Broad teaching and industrial experience in New York-New Jersey area. Transportation consultant. Formerly on the staff of New York Academy of Advanced Traffic. Interested in teaching/research in Boston area. Available in June or September, 1970. E2138

Micro, industrial organization statistics: Man, 32, married; A.B., M.A., all requirements for Ph.D. completed except for dissertation, dissertation completion is not imminent. Currently assistant professor at college in Northeast. Seven years of college teaching experience with active participation in faculty matters. Seeks research and/or teaching position. E2139

Econometrics, micro- and macroeconomic theory, public finance, industrial organization: Man, married; recent Ph.D. State University of New York Research Foundation Fellowship. Three years of college teaching; presently assistant professor at State University of New York. Desires teaching and/or research position. Available in September, 1970 or 1971. E2140

Public finance, comparative economic systems, history of economic thought: Man, 27; Ph.D. Five years of college teaching; several publications. Desires teaching position with significant research opportunities. Would consider department chairmanship or directorship of center involving economics education or public finance. Available in June or September, 1970. E2141

Principles, economic development, natural resources economics, regional economics, international economics, Latin American economics, public finance, demography: Man, 36, married; A.B., M.A., Ph.D. dissertation in progress. Honors and awards. Two years of research in Latin America. Five years of diversified college teaching. Fluent in Spanish. Seeks permanent teaching position in challenging undergraduate program. Available in June or September, 1970. E2142

Principles, money and banking, international economics, economic doctrine, theory: Man, 57, married; Ph.D., Wisconsin. Twenty-two years of teaching experience at the college level. Experience as chairman of economics department in liberal arts college; extensive travel in Europe; linguistic ability. Wishes to relocate at a liberal arts college or take foreign assignment. Available in September, 1970. E2143

Political economy of technological change, economic history, microeconomics, government and business: Man, 58; Ph.D. Twenty years of teaching graduate and undergraduate; sixteen years of research and administration in government and nonprofit organizations. Substantial book and journal publications. Desires college or university teaching or administrative position. E2144

Regional economics, urban economics, economics of construction industry: Man, in 40's, married; B.S. (civil engineering), M.B.A., Ph.D. candidate. Registered professional engineer and land surveyor. Experience includes engineering, construction management, teaching, and administration. Desires consulting, industrial, governmental, or teaching position where total range of experience can be applied to solving current problems. E2145

Economic theory, economic development, labor economics, economic thought: Man, 29, married; M.A., Ph.D. in progress. Eight years of university teaching in India and U.S. Modest list of publications. Seeks teaching position in college or university. Résumé on request. Available in June or August, 1970. E2146

Money and banking, international finance, public finance, economic theory and thought, economics of developing areas: Man, 36, married; U.S. permanent resident; M.A., Ph.D. expected in summer, 1970 (finance with concentration in the monetary area), both degrees from a major highly reputable university. Experience includes six years of college teaching in the U.S. and four years of banking abroad. Excellent references. Seeks a teaching position in U.S. college or university; no geographic restriction. Available in September, 1970. E2147

Macroeconomics, monetary economics, labor: Man, 31, married; course requirements for Ph.D. complete. Five years of teaching experience at large state university. Interested primarily in eastern or southern institution interested in quality teaching. Has taught variety of courses, including both intermediate micro and macro. Prefers macro, monetary, and labor but would willingly fill in where needed. Not interested in "publish or perish" situation but one stressing instruction, student contact, departmental duties, etc. Would prefer location in or near major metropolitan area but not crucial. Available in fall, 1970, if conditions look favorable. E2148

McGRAW-HILL

is pleased to announce the new 8th edition of
Samuelson: **ECONOMICS** . . . now available.

In the author's words ". . . this is the most complete revision since the 4th edition. RELEVANCE is the goal; complacent, Establishment economics the enemy. . . ."

Two new chapters deal with:

- *Economics of the poverty program and inequality
- *Economics of race and urban problems
- *The quality of economic life and environmental pollution
- *The military-industrial complex and exploitation

Beyond updating material to the 1970's, the 8th edition introduces:

- A restructuring of microeconomics, stressing economics of welfare
- A restructuring of macroeconomics, emphasizing the role of money supply in income determination

New topics that are treated in greater depth than in other textbooks are:

- | | |
|--|--|
| *Monetarism (Chicago and other versions) | *Galbraith's New Industrial State |
| *New Left and Imperialism | *Negative Income Tax |
| *Conglomerates | *Flexible Exchange Pegs and Paper Gold |
| *Volunteer Army | |

Also, new and revised supplements:

- **READINGS IN ECONOMICS**, Sixth Edition, Paul A. Samuelson
- **STUDY GUIDE**, Romney Robinson (University of Toronto)
- **INSTRUCTOR'S MANUAL**, Paul A. Samuelson, Maureen Crandall (Wellesley)
- **TEST BANK**, Paul A. Samuelson, Maureen Crandall (Wellesley)
- **ECONOMICS: An Introductory Program**, Edward Foster (University of Minnesota)
- **VISUAL AIDS**: 8mm film loops are available from the Visual Learning Corporation; Overhead Transparencies are available from the publisher

ECONOMICS: Principles, Problems, and Policies, Fourth Edition

Campbell R. McConnell, University of Nebraska. 815 pages, \$9.95

This is the latest edition of a leading text that introduces all the principles essential to an understanding of fundamental economic problems and the policy alternatives society may utilize to contend with these problems. The fourth edition of McConnell reflects new developments in economic analysis and changes in economic institutions, problems, and policies. In addition, the fourth edition represents a substantial refinement in the pedagogical effectiveness of the text. Major changes occur in three specific areas: international economics, economic growth, and fiscal policy. The first includes a thorough theoretical development of comparative advantage, an improved discussion of the balance of payments and international finance, and a general updating of the discussion of international trade problems and policies. Economic growth stresses the importance of growth, provides an analytical framework and reorganizes the subsequent material into a more systematic discussion of the growth record of the United States and growth-related problems. Fiscal policy has been extensively reworked to include new concepts of theory and policy: e.g., the CEA's "budget line" analysis is used to distinguish between automatic and discretionary fiscal policy and to explain full-employment surplus, fiscal drag, the fiscal dividend, etc.

Various supplementary aids are available: Test File, Instructor's Manual, Economic Issues: Readings and Cases, Third Edition, Study Guide, Economic Concepts: A Programmed Approach, Program and Study Guide, Overhead Transparencies, and Transparency Masters.

AN AXIOMATIC SET THEORETIC APPROACH TO MATHEMATICAL ECONOMICS:

INTRODUCTION TO CONTEMPORARY MICROECONOMICS

Vivian Walsh, University of Washington. 352 pages, \$5.95

After presenting a general theory of choice from micro-theory, the author then develops models of consumption, production, and exchange. Elementary notions of first order logic, of naive set theory, and of certain topological properties of point sets, are developed *within* the exposition. Set theoretic concepts are utilized throughout to show the novice the place of economic theory at the core of the vastly developing new science of optimal decision making. No familiarity with mathematics is assumed.

THREE ESSAYS ON THE STATE OF ECONOMIC SCIENCE

Tjalling C. Koopmans, Yale University. 231 pages, \$7.95, \$3.50 (soft-cover)

The author, one of the pioneers in explaining the new set theoretic in mathematical economics, examines different phases of economic science in these three essays. The first essay discusses the descriptive theory of competitive equilibrium and the normative theory of the use of prices for the efficient allocation of resources; the second explores the character and basis of economic knowledge; and the third analyzes four tool developments and investigates their implications for future work in theory and empirical investigations.

INTRODUCTION TO GENERAL EQUILIBRIUM THEORY AND WELFARE ECONOMICS

James Quirk and Rubin Saposnik, both of the University of Kansas. *Economics Handbooks Series*. 240 pages, \$8.95

This highly acclaimed book examines relatively recent developments in general equilibrium analysis without utilizing the rather advanced mathematics that have been a tool of such developments over the past 15 years. Various topics such as existence and stability of competitive equilibrium, welfare economics, etc., are discussed against the pervasive theme of general equilibrium.

THE LOGIC OF THE PRICE SYSTEM

Paul B. Trescott, Southern Methodist University. Available Spring, 1970

Utilizing a minimum of calculus, this book stresses the application of microeconomics to levels beyond the individual firm or consumer. The author elaborates on such economy-wide matters as factor pricing and income distribution, resource allocation, international trade and economic growth. The logic and limitations of government action in a market economy are also discussed.

ECONOMICS OF INTERNATIONAL TRADE

Richard I. Leighton, State University of New York at Binghamton. 240 pages, \$8.95

This short, theoretical book stresses both pure theory and monetary economics. The author examines the economic variables that determine the level of international commerce and their interaction with the domestic economy. With a basically post-Keynesian approach, this book uses both graphs and simple algebra to explain the theory. A major feature is the extensive treatment given to consequences of capital flow.

ECONOMICS AS A SCIENCE

Kenneth E. Boulding, University of Colorado. In Press

This world renowned author provides introductory economics students with a perspective of the scope and setting of economics and the place which it occupies in the intellectual community. Each chapter covers a particular aspect of economics in relation to some larger whole; e.g., economics as a social science, an ecological science, a political science, a mathematical science, etc.

CONTEMPORARY ECONOMICS: Principles and Issues

Leonard S. Silk, The Brookings Institution. 500 pages, \$8.95

Analytical and problem-oriented in its approach, this text presents many real-world examples that are within the experience of the average student. The scope includes detailed treatments of major contemporary problems (e.g., poverty, inflation, unemployment, etc.) and illustrations of how fundamental economic concepts can be applied to clarify and help solve these problems. *Readings in Contemporary Economics* by Leonard Silk, is also available, as well as a Study Guide by Arthur L. Welsh, Joint Council on Economic Education, an Instructor's Manual, and a Test File.

INDUSTRIALIZATION AND ECONOMIC HISTORY: Theses and Conjectures

Jonathan R. I. Hughes, Northwestern University. Available Summer, 1970

This book's general purpose is to examine the economic, social and political effects that industrialization has had in certain countries since the 18th century. Specifically, it discusses, among other topics, the consequences of war as conducted by industrial nations; explains why modern economic and political conflicts are the outcome of industrial history; and traces the modern international system as a "descendant" of the pre-1914 Gold Standard.

BASIC ECONOMIC PRINCIPLES, Second Edition

Robert A. Lynn, Kansas State University. 384 pages, \$8.50

The purpose of this one-term text is to: (1) present a survey of economic analysis, issues, and institutions; (2) relate this material to the important economic issues that are currently confronting the United States; and (3) familiarize the student with the basic tools that economists use to analyze economic situations. Two of the major considerations of the book are underutilization of resources, and the inability of the American economy to deal with its massive poverty problem. A Study Guide, Instructor's Manual, and Test File are also available.

ARBITRATION AND COLLECTIVE BARGAINING

Paul Prasow and Edward Peters, both of the University of California, Los Angeles. McGraw-Hill Series in Management. 480 pages, \$12.95

The thesis of this book is that the dynamics of collective bargaining can best be understood by studying arbitral issues and concepts. The authors, therefore, have provided a theoretical framework for examining the arbitration process and, in so doing, explore a wide range of substantive and procedural issues. Analytical concepts, rather than techniques, are stressed and a variety of actual case studies are included.

PRODUCER, CONSUMER, AND SOCIAL CHOICE

O. H. Brownlee and John A. Buttrick, both of the University of Minnesota.
352 pages, \$9.95

This book applies a unified theory to problems of producer, consumer, and social choice under conditions of uncertainty as well as certainty. Production relationships of the type employed in linear programming are presented at a reasonable level of mathematical difficulty to teach applied decision problem-solving.

FUNDAMENTAL METHODS OF MATHEMATICAL ECONOMICS

Alpha C. Chiang, University of Connecticut. 576 pages, \$12.50

Requiring only a background in high school algebra, this modern economics text (1) carefully develops mathematical techniques (e.g., set theory, convex sets, etc.) and then fits them into an economic framework; (2) interweaves illustrations and applications of economic concepts and utilizes "models" of equilibrium analysis, comparative-static analysis, dynamic analysis, optimization problems, linear programming, and game theory. In many instances, intuitive explanations are provided in addition to formal derivations.

MAN AND ECONOMICS

Robert A. Mundell, University of Chicago. 208 pages, \$5.95, \$3.50 (soft-cover)

The purpose of this book is to acquaint the reader with the tremendous power of economics as a system of thought and as a guide to the solution of problems of choice in all aspects of life. It is a simple, straightforward statement of the principles of economics, conceived of as the science of choice, and the application of these principles to problems of economic policy and world order.

REGIONAL ECONOMICS

Hugh O. Nourse, University of Illinois. *Economics Handbooks Series.*
256 pages, \$9.50

Theoretical in its approach, this text examines those economic theories which help to explain: city locations and hierarchies, industrial location patterns, land use patterns, the short-run impact of industrial change upon employment in the community, and the long-run differentials of income per capita between regions.

ECONOMIC DECISION MODELS FOR ENGINEERS AND MANAGERS

James L. Riggs, Oregon State University. 401 pages, \$11.95

Presents a concise, lucid, and broad coverage of quantitative methods for economic decision making. Topics are carefully selected and organized to integrate the important decision making tools, progressing from models assuming certainty to those assuming uncertainty. In addition to its emphasis on economic aspects and its inclusion of probabilistic considerations, this text combines deterministic decision models with the usual engineering economy subjects.

McGraw-Hill Book Company • 330 W. 42nd Street • New York, New York 10036

International — for a new selection in 1970

BASIC ECONOMICS

By Micha Glsser, The University of New Mexico; and
Peter S. Barth, Ohio State University

This is a concise treatment of basic economic analysis and policies that features the inclusion of recent developments in contemporary research. Most of the chapters end with a short summary of ideas and questions that have been challenging the profession in recent years. The presentation is analytical but not at the expense of policy. 1970, 455 pp., \$8.00

THE ECONOMIC AND FINANCIAL SYSTEM

By Richard A. Ward, University of Southern California

A presentation of macroeconomics, this book emphasizes an integrated treatment of income flows and financial assets and develops the integrated treatment of the mutual relationship between claim creation and product demand. 1970, 302 pp., \$8.50

INTRODUCTION TO PRICE THEORY, 2D ED.

By Micha Glsser, The University of New Mexico

In this book the basic elements of price theory are presented by using simple tools such as tables and diagrams. Additions in this *Second Edition* include the theory of capital budgeting, general equilibrium and welfare economics, and linear programming. 1969, 413 pp., \$7.95

MANPOWER PROBLEMS AND POLICIES: FULL EMPLOYMENT AND OPPORTUNITY FOR ALL

Edited by John A. Delehanty, Kansas State University

The major theme of the book is problems and policies involved in moving beyond overall full employment to achieve opportunity for all. Among other topics, the readings deal with the lack of opportunity in urban and rural areas. 1969, 412 pp. (paper), \$4.95

MONEY, BANKING, AND INCOME: THEORY AND POLICY

By Pham Chung, The University of New Mexico

This is an integrated treatment of the historical, institutional, and theoretical aspects of money and banking. Current international financial problems such as gold outflows and the U.S. dollar, the need for reform of the international monetary system, and recent reform plans are covered. March, 1970

THE POLITICAL ECONOMY OF COMMUNIST CHINA

By Jan S. Prybyla, The Pennsylvania State University

A survey of the development of China's communist economy from 1949 through 1969. It covers the social, political, and cultural philosophy and policies of the communist regime. June, 1970

TARGETS AND INDICATORS OF MONETARY POLICY

Edited by Karl Brunner, Ohio State University

Most of the papers in this volume were presented at a conference in 1966 sponsored by the American Bankers Association and the Institute of Government and Public Affairs, UCLA. The focus is on the state of knowledge of monetary behavior and its relation to actual monetary policy. 1969, 352 pp., \$8.50, A Chandler Publication

Examination copies are available upon request. Write to:

International Textbook Company

An Intext Publisher

Scranton, Pennsylvania 18515

The Function and Analysis of Capital Market Rates—James C. Van Horne, Graduate School of Business, Stanford University. Focusing on market rates of interest, this new text develops a conceptual framework for analyzing interest rates. Synthesizing recent theoretical and empirical literature, the text analyzes the role of interest rates in channelling savings in an economy as well as differences in yield among financial instruments. The author considers why financial instruments exist in an economy, the efficiency of financial markets, the measurement of savings flows, and financial equilibrium brought about by interest rates. The concentrated focus of this study is on interest rates in a broad sense (including equity instruments) under conditions of uncertainty. February 1970, approx. 192 pp., \$3.95 (33193-4)

Capital Markets and Institutions, 2nd Edition, 1970—Herbert C. Dougall, Stanford University—"The aim of the second edition is to present a careful but uncomplicated study of the institutions through which long-term funds are funneled into several markets . . . so that the resulting yields can be noted and analyzed. The data are again integrated in a master table for the capital market as a whole. All data have been updated through 1968 and together with partial information for 1969 form the basis for appropriate comment."—from the preface

foundations of finance

Forthcomings:

Portfolio Analysis, Stephen H. Archer and Clark Francis

Published books in the Series

Capital Markets and Institutions, Herbert E. Dougall

Accounting Flows: Income, Funds and Cash, Robert K. Jaedicke, and Robert T. Sprouse

Investment Decisions and the Capital Costs, James T. S. Porterfield

Optimal Financing Decisions, Alexander Robichek and Stewart C. Myers

The Scope and Methodology of Finance, Fred J. Weston

**for further information write box 903,
Prentice-Hall, englewood cliffs, n.j. 07632**

The Economic Problem, Second Edition—Robert L. Helbroner, The New School for Social Research. The Second Edition of this well-known text presents an introductory examination of economics focusing on three principal areas: a broad understanding of economic history; the basic elements of macroeconomics, stressing intersectoral relationships; and the essentials of microeconomics and foreign trade. Dr. Helbroner has made the book more analytical, especially in the areas dealing with the concept of equilibrium, the concept of GNP, and stocks and flows. In chapter 16, the expansion of money is clearly presented through the use of T accounts. A revised chapter, on money and the macro-system,

2nd edition

Introduces a new presentation on money and inflation. Chapter 19, on the problem of growth, now includes a new analytical section on capital output. Chapter 29 provides new material on the gold crisis and a discussion of the American balance of payments in 1968. In chapter 30, the material on USSR planning has been rewritten and updated to include the new reforms. The Second Edition contains revised diagrams, many utilizing color. An expanded teacher's manual, a companion readings book, and a newsletter are also available. February 1970, approx. 672 pp., \$9.50 (22698-5)

for further information write box 903

Key Features of the Second Edition

readability—reviews indicate this is a book students enjoy reading, pick up and read again, and upon completion understand, appreciate, and retain. Excessive detail not appropriate to an introductory course is omitted.

historical material—Parts I (8 chapters) and V (3 chapters) are primarily historical, which leads to a real understanding of the economic problem through the impact of historical factors. The student begins thinking about problems in economic terms.

the economic problem

Supplements for The Economic Problem include:

Student Guide, by John Hannaford, Ball State University. Specifically designed to keep the student's focus on the essentials developed in the text. Summaries contain the substance of each chapter. Vocabulary and conceptual understanding are tested. The programmed technique is used throughout to help teach difficult concepts.

A separate book of readings edited Arthur MacEwen and Thomas E. Weisskopf (both of Harvard University) will be available for use with the Second Edition.

Teacher's Manual, by John Hannaford, Ball State University. Contains a reproduction of the text with marginal notes designed to help the teacher, and complete testing materials.

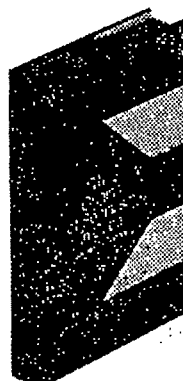
Free Newsletter, by Robert Hellbroner. Focuses on economic issues and problems.

Prentice-Hall, englewood cliffs, n.j. 07632



Cornell University Press

- **CASTRO'S ECONOMY**
- **OPPORTUNITY COST**
- **AGRICULTURAL GROWTH**



The Origins of Socialism in Cuba

By JAMES O'CONNOR. Sure to arouse controversy, this book contends that Cuban socialism developed organically out of the traditions and long-standing needs of the people. The author maintains that socialism was inevitable if Cuba was to be rescued from economic stagnation and political corruption. He supports his thesis with a close examination of the Cuban economy before the revolution and of the steps Castro took to initiate socialism.

344 pages, tables. \$10.00

Information for Management Decisions

A System for Economic Analysis and Accounting Procedures

By YAAQOV GOLDSCHMIDT. The author describes his system for analyzing a business operation in order to provide management with better information for reaching decisions. Called the operational accounting and analysis system (OAAS), it is based both on economic theory and on accounting procedures. A comprehensive guide to data necessary for decision-making and control.

376 pages, figures, tables. \$11.00

NOW IN PAPERBACK

The Economics of Agricultural Development

By JOHN W. MELLOR. "This book fills an important gap in the economic development literature."—*The American Economic Review*

"This is an important book reflecting the author's extensive experience ... in the underdeveloped world ... It is meticulous, systematic, and a remarkably comprehensive analysis."

—*Journal of Agricultural Economics*
416 pages, 6 illustrations. \$3.45

Available in cloth, \$10.00

Cornell University Press
ITHACA AND LONDON

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

3
pages
to help
you
reevaluate
your

Economics Texts for 1970

The popular
alternative to the big,
introductory text—

INTRODUCTION TO ECONOMICS SERIES

Edited by KENYON A. KNOFF, Whitman College.

Used at over 500 colleges and universities

Now there is a new book in the series

THE ECONOMICS OF THE PUBLIC SECTOR

By ROBERT H. HAVEMAN, Grinnell College.

Here is a current guide to economic choice in the governmental sector of the economy. It explores the efficiency and the income redistributive effects of governmental decisions. It discusses such pertinent questions as why we spend more money on missiles than cities. The material is unified by the concept of the Principle of Maximum Social Gain.

1970

Approx. 320 pages

\$6.95 (Cloth)

\$2.95 (Paper)

This series has been so successful that now
there are 4 revised and updated editions—

THE ECONOMICS OF POVERTY

Second Edition

By ALAN B. BATCHELDER, Kenyon College.

1970 In press

NATIONAL INCOME AND EMPLOYMENT ANALYSIS

Second Edition

By ARNOLD COLLERY, Amherst College.

1970 Approx. 192 pages \$5.95 (Cloth)
\$2.95 (Paper)

INTERNATIONAL ECONOMIC PROBLEMS

Second Edition

By JAMES C. INGRAM, University of North Carolina, Chapel Hill.

1970 Approx. 224 pages \$6.95 (Cloth)
\$2.95 (Paper)

THE MARKET SYSTEM

Second Edition

By ROBERT H. HAVEMAN, Grinnell College, and
KENYON A. KNOFF.

1970 Approx. 320 pages \$6.95 (Cloth)
\$2.95 (Paper)

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Economics

Texts for 1970

MONEY AND BANKING:

An Introduction to Analysis and Policy Second Edition

By JOHN G. RANLETT,
Sacramento State College.

Dealing with the institutions of our contemporary money and banking system, the second edition integrates the treatment of domestic money and banking, theory, and policy, adds new IS and LM derivatives, plus new material on the Money Supply Theory. An Instructor's Manual is available.

1969 543 pages \$9.95

ECONOMETRICS

By RONALD J. WONNACOTT and THOMAS H. WONNACOTT, both of the University of Western Ontario.

This new text for 1970 conveys the notion of structure in econometrics. The first half of the text builds on the later chapters of *INTRODUCTORY STATISTICS*, by WONNACOTT & WONNACOTT (Wiley, 1969), and provides an elementary introduction to the concepts used in econometrics. The second half of the text covers the same topics on a more sophisticated and in-depth level, using linear and matrix algebra.

1970 Approx. 480 pages \$12.95

THE COURSE OF AMERICAN ECONOMIC GROWTH AND DEVELOPMENT

By LOUIS M. HACKER, Columbia University.

Here is an introduction to the factors and forces responsible for economic growth and the resulting social and institutional development in the United States.

Wiley Series in American Economic History, edited by Ralph Andreano, the University of Wisconsin.

1970 Approx. 368 pages Cloth: \$8.95
Paper: \$4.95

MONEY AND FINANCE:

Readings in Theory, Policy and Institutions Second Edition

Edited by C. D. CARSON, Columbia University.
In preparation: available in 1970.

INTRODUCTION TO ECONOMIC STATISTICS

By WILLIAM C. MERRILL and KARL A. FOX,
both of Iowa State University.

Statistical techniques are related to basic probability; the importance of assumptions made in the development of statistical models is demonstrated; and attention is focused on descriptive statistics, index numbers, and time series analysis. A Teachers Manual is available.

1970 658 pages \$11.95

LABOR ECONOMICS:

The Changing Labor Scene

By C. GLYN WILLIAMS, University of South Carolina.

The topics in this text have been selected on the basis of their importance to an understanding of economic developments in the contemporary American labor scene—from the supply of and demand for labor to unions, management, and the economy in general.

1970 Approx. 576 pages \$9.95

READINGS IN MICROECONOMICS

Edited by DAVID R. KAMERSHEN,
University of Missouri.

Selected readings in price theory emphasize analysis rather than description. Articles containing mathematics or statistics have enough economics in them so that they can be usefully employed in intermediate and graduate courses in economics.

1969 607 pages Cloth: \$9.95 Paper: \$6.95

ECONOMIC DEVELOPMENT: The Cultural Context

By THOMAS R. DeGREGORI, University of Houston; and ORIOL PI-SUNYER, University of Massachusetts.

This is an interdisciplinary approach to the investigation of economic, social, and technological change in traditional and primitive societies. The authors make use of analytical and illustrative references drawn from the experience of advanced industrial societies.

1969 138 pages Cloth: \$6.95 Paper: \$3.50

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Economics Texts for 1970

SALES ANTICIPATIONS AND INVENTORY BEHAVIOR

By ALBERT A. HIRSCH, U.S. Department of Commerce; and MICHAEL C. LOVELL, Carnegie-Mellon University.

This new text provides empirical evidence on the way businessmen form their expectations, on the accuracy of their predictions of future sales volume, and on the impact of forecasting errors on inventories and production. The book focuses on new information provided by the Manufacturers' Inventory and Sales Expectations Survey of the Department of Commerce.

1969 256 pages \$10.95

MICROECONOMIC ANALYSIS

By ALLAN J. BRAFF, University of New Hampshire.

"... written to make the reader aware of the intellectual substance and excitement of microeconomic analysis... contains extended discussions of imperfect markets, the problem of identification, stability conditions, the process of adjustment to equilibrium, the integration of classical utility theory with indifference curve analysis, behavioral theories of the firm, general equilibrium, and welfare economics."—from the Preface

1969 295 pages \$7.95

INTERMEDIATE MACROECONOMIC THEORY

By JOHN LINDAUER, Claremont Men's College and the Claremont Graduate School.

A graphical model of the macroeconomic aspects of an economy is developed and used in this text to analyze the various macroeconomic conditions and problems that might arise. Aggregate supply and demand are examined as well as growth, inflation, unemployment, and monetary and fiscal policies.

1968 334 pages \$8.95

AN INTRODUCTION TO PUBLIC FINANCE

By HERBERT E. NEWMAN, Hood College.

This survey of public finance in terms of the United States' political and institutional limitations covers problems at all levels—from the economic effects of government finance decisions to the political and administrative environment within which government finance decisions are made.

1968 677 pages \$9.95

LABOR THEORY

By RICHARD PERLMAN, University of Wisconsin.

This well balanced, advanced treatment of the theoretical basis of major topics of labor economics features discussions of the Phillips curve and the inflation-full employment trade-off, the effect of income-maintenance plans on work incentive, and the structural-inadequate demand controversy.

1969 237 pages \$8.95

INTRODUCTORY STATISTICS

By THOMAS H. WONNACOTT and RONALD J. WONNACOTT, both of the University of Western Ontario.

This text introduces statistical concepts on a sophisticated level, stresses inferential statistics, and gives an excellent introduction to decision theory—all with clarity and a minimum of mathematical complexity.

1969 403 pages \$9.95

WHAT IS ECONOMICS?

By JOHN E. MAHER, Joint Council on Economic Education, and Director of Developmental Economic Education Program (DEEP). Here is a concise answer plus further questions that give the reader a guide toward understanding the discipline.

1969 174 pages Cloth: \$5.95 Paper: \$2.95

wiley

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

JOHN WILEY & SONS, Inc.
605 Third Avenue, New York, N.Y. 10016
In Canada: 22 Worcester Road, Rexdale, Ontario

THE HARRIED LEISURE CLASS

Staffan B. Linder

"I greatly recommend The Harried Leisure Class. It combines precise economic analysis with good writing and brilliant insight."—John Kenneth Galbraith.

192 pages

SBN 231-03302-8

\$7.00

BANKERS' DIPLOMACY:

Monetary Stabilization in the Twenties

Richard H. Meyer, Federal Reserve Bank of N.Y.

A careful inquiry into the impact of the negotiation of international monetary stabilization loans and credits during the years 1926-1928 on the policies of cooperation among the world's central banks. Mr. Meyer points out similarities to today's international currency crisis in an effort to provide a reevaluation of previous experience which might help avoid a recurrence of disaster.

184 pages

SBN 231-03325-7

\$8.00

POVERTY AND PLENTY ON THE TURKISH FARM:

An Economic Study of Turkish Agriculture in the 1950's

Eva Hirsch, Rutgers University

Through an intensive analysis of several villages, Dr. Hirsch pieced together this picture of Turkish agriculture in the 1950's, which not only throws much light on the topic of income distribution but also provides statistical and analytical methods which can be utilized in the study of other underdeveloped countries.

Modern Middle East Series, I, Middle East Institute, Columbia University.

\$6.00

LABOR AND DEVELOPMENT IN LATIN AMERICA

Joseph R. Ramos, University of Chile

This study analyzes the role of the labor force in Latin American development since World War II. The author inquires into shifts in the composition of the labor force and the effects of the importation of modern technology.

Institute of Latin American Studies, Columbia University.

320 pages

SBN 231-03250-1

August \$12.50

CRISIS IN BRAZIL

Octavio Ianni, University of Sao Paulo

Viewing the 1964 coup d'etat as the watershed between the era of popular democracy and the present stage of dependency relations in its capitalistic system, Octavio Ianni presents an exhaustive and illuminating economic, political, and sociological study of the dilemma which Brazilian society continues to face: "dependence and capitalism, or independence and socialism."

224 pages

SBN 231-03221-8

\$7.50



Columbia University Press

ADDRESS FOR ORDERS: 136 So. Broadway
Irvington, N.Y. 10533

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers



Leading Issues in Economic Development:

STUDIES IN INTERNATIONAL POVERTY *SECOND EDITION*

By GERALD M. MEIER, *Stanford University*. This new type of course book, extensively revised and expanded in the Second Edition, emphasizes strategic policy issues in accelerating the development of poor countries. The leading issues in this new edition are now less conceptual and more policy-oriented—less a function of academic debates among economists and more the outcome of perplexing experiences of development policy during the past two decades. The book examines the issues of dualistic development, mobilization of domestic and foreign resources, allocation of investment resources, industrialization and agriculture, foreign trade and development, human-resource development, and development planning. A substantial portion of the work consists of commentaries by Professor Meier, in the form of supplementary text Notes that integrate and clarify the diversified materials.

1970

800 pp.

\$9.95

Society and Economic Growth:

A BEHAVIORAL PERSPECTIVE OF SOCIAL CHANGE

By JOHN H. KUNKEL, *Arizona State University*. This work examines the inner workings and some of the problems of social change and economic development. Drawing on research in sociology, anthropology, economics, and psychology, it covers such topics as social structure, government, culture, and systems, and presents a significant contribution to the literature of social and economic dynamics.

1970

400 pp.

\$6.50

Growth Theory:

AN EXPOSITION

By ROBERT M. SOLOW, *Massachusetts Institute of Technology*. The six lectures in this volume present an introductory survey of the modern macroeconomic theory of growth and its application to economic policy. Included in the book are discussions of new research topics, such as growth in a monetary economy, and optimal programs of capital accumulation.

1970

148 pp.

cloth \$4.50

paper \$1.95

Power and Poverty:

THEORY AND PRACTICE

By PETER BACHRACH, *Temple University*, and MORTON S. BARATZ, *Boston University*. A significant contribution to the field of community power studies, this book presents a new approach to the analysis of power and the correlates of power: authority, influence, and force. Part I consists of a theoretical analysis of the political process, with particular reference to the concept of non-decision-making, culminating in a presentation of a model of the political system. Part II applies the theoretical model to the relationship between the anti-poverty effort and the political process in Baltimore, Maryland.

1970

240 pp.

cloth \$6.00

paper \$2.50

OXFORD UNIVERSITY PRESS

200 Madison Avenue, New York, N.Y. 10016

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Oxford University Press



The Theory of Economic History

By SIR JOHN HICKS, *Oxford University*. Economists are inclined to think of the market economy as always existing, just developing, or "growing"; historians know very well that this is not the case. This book builds a bridge between these two different approaches. Its subject is the evolution of the market economy, its forms and institutions. "Another particularly brilliant chapter devoted to the growth of money and credit makes clear what many historians shroud in impenetrable detail . . . exciting reading."—*The Economist* Cloth, \$5.00. A *Galaxy Book*, GB 302, paper, \$1.95

Modernizing Peasant Societies

A COMPARATIVE STUDY IN ASIA AND AFRICA

By GUY HUNTER. "One of the best balanced treatments on the modernization of the peasantry yet to appear. . . . A learned book that should be read by policymakers in the new nations no less than by scholars in the old nations."—IRVING LOUIS HOROWITZ, Rutgers University. "A significant study—focusing on two major problems of economic development: the agricultural bottleneck and the transfer of institutions (as well as ideas and technology) from richer to poorer countries."—GERALD M. MEIER, Stanford University. (*Institute of Race Relations*) Cloth, \$6.50. Paper, \$2.50

The Colonial Heritage of Latin America

ESSAYS ON ECONOMIC DEPENDENCE IN PERSPECTIVE

By STANLEY J. STEIN and BARBARA H. STEIN, both of Princeton University. "An extremely skillful synthesis. . . . Nowhere else can one find such a penetrating analysis of Latin America's economic ties with Europe. I have no doubt that this persuasive analysis will immediately become required reading in Latin American history courses everywhere."—THOMAS E. SKIDMORE, University of Wisconsin Cloth, \$5.00. Paper, \$1.50

Empire to Welfare State

ENGLISH HISTORY 1906-1967

By T. O. LLOYD, *University of Toronto*. Professor Lloyd's history of Great Britain in the twentieth century is distinctive in its focus upon British economic history and current economic policy, rather than upon political and diplomatic history alone. While the volume offers a highly comprehensive survey of the economic trends of the period, it also gives special attention to social problems and the history of ideas. It is the first volume of *The Short Oxford History of the Modern World*.

Text figures, maps, Cloth, \$10.00. Paper, \$2.95

OXFORD UNIVERSITY PRESS

200 Madison Avenue, New York, N.Y. 10016

Please mention *THE AMERICAN ECONOMIC REVIEW* When Writing to Advertisers

Oxford University Press



Theory of Economic Growth

By MICHIO MORISHIMA, *University of Essex*. In this book the author discusses economic growth with the framework of the theory of general equilibrium in the attempt to resolve much of the controversy on growth due to differences of assumptions made by various authors. He also makes a contribution to the theory of dynamic economics. \$12.80

E. D. Morel's History of the Congo Reform Movement

Edited by W. R. LOUIS, *Yale University*, and J. STENGERS, *University of Brussels*. E. D. Morel, as one of England's most powerful reformers, denounced the "rubber atrocities" of King Leopold's Congo and stirred the humanitarian conscience of the English public. Under his leadership the Congo reform movement figured prominently in the political history of England, Europe, and Africa in the decade prior to the first World War. This volume is Morel's previously unpublished account of his discovery of the exploitation involved in the Congo and of the events leading to the founding of the Association in 1904. \$6.75

The Economics of Money

SECOND EDITION

By A. C. L. Day, *University of London*. This is a concise introduction to the theory and practice of the financial system, revised by the author of this OPUS edition. It is intended for both the intelligent general reader who wishes to understand something of economic and financial matters, and for the student. It can be regarded as an introduction to Mr. Day's longer work, *AN OUTLINE OF MONETARY ECONOMICS*. An OPUS Book, No. 31, paper, \$1.05

Contemporary China, 1962-1964

VOLUME VI

Edited by E. STUART KIRBY, *University of Aston*. Volume VI covers a period of many important developments, in particular the marked economic improvements of 1962, which were strongly sustained and widened in the following years. The period also saw more direct forms of confrontation between the Chinese and Russian Communist parties. This volume consists of documentary material dealing with specific current problems present in China during this era.

(*Journal of Oriental Studies*)

OXFORD UNIVERSITY PRESS

200 Madison Avenue, New York, N.Y. 10016

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Addison Wesley Texts:

entering a new era in economics

Introduction to Statistical Methods

by Donald L. Harnett, *Indiana University*

This text is intended for an introductory course in statistics for students of business and the social sciences at the sophomore or junior level, or for first year graduate students. Prerequisites include a course in college algebra and an exposure to the elements of calculus.

In press (1970)

Choice Against Chance: An Introduction to Statistical Decision Theory

by John Althison, *University of Glasgow*

This book provides a readable, elementary introduction to decision-making under uncertainty. The relationship and correspondence of mathematical models to reality are stressed, leading students and specialists in many fields to see the wide variety of potential applications of statistical decision theory.

In press (1970)

Managerial Economics: Decision Making under Certainty for Business and Engineering

by Floyd E. Gillis, Jr., *Purdue University*

This is the first book to come to grips with the problem of applying economic theory to business decisions on the basis of conditions as they are today.

296 pp, 71 illus, \$9.50 (1969)

Macroeconomic Analysis: An Introduction to Comparative Statistics and Dynamics

by Thomas Dernburg, and Judith Dukler Dernburg, *Oberlin College*

This book attempts to push the student's knowledge of macroeconomics beyond the first-course level, and it develops those mathematical tools with which most of the students taking such a course are not familiar.

292 pp, 55 illus (1969)

Economics

by Bernt P. Stigum and Marcia L. Stigum, *Northwestern University*

Designed for introductory college courses in economics, this book gives the student the analytic tools and institutional background necessary to understand economic problems and to evaluate economic policies.

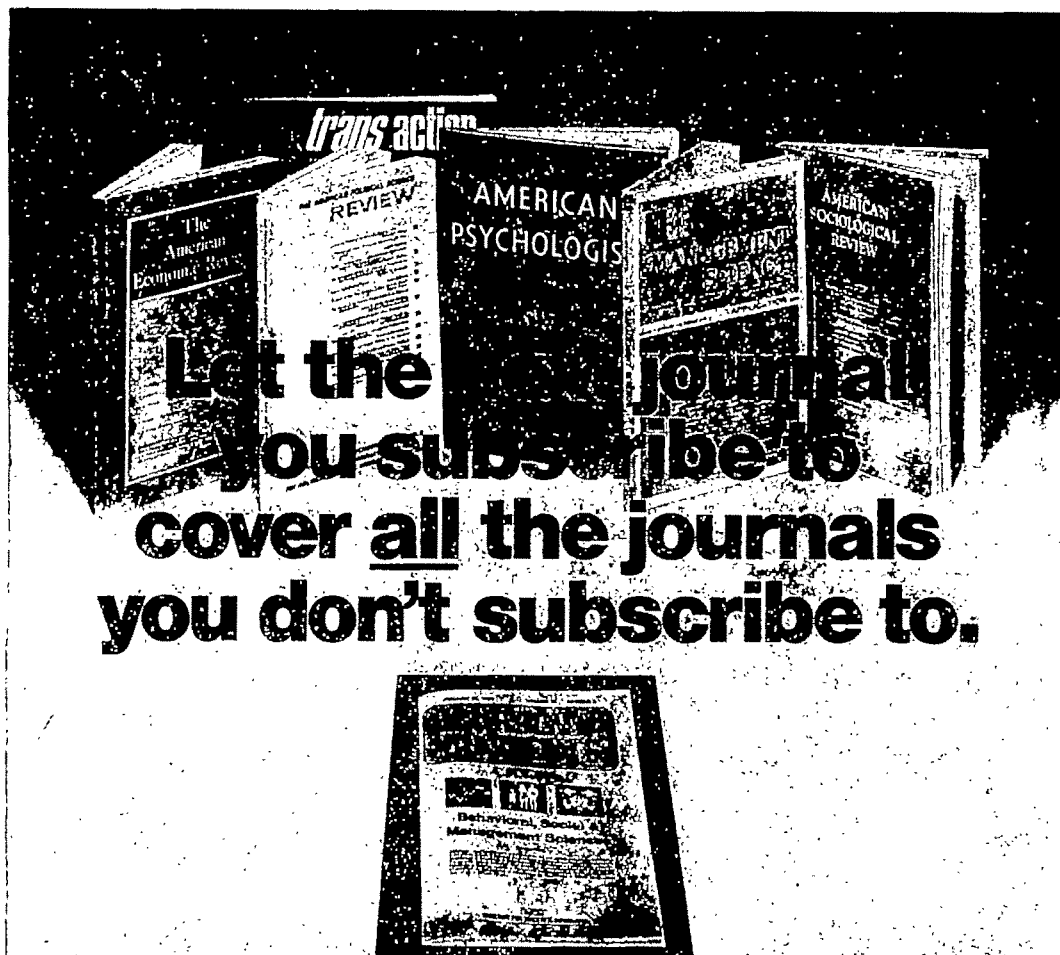
793 pp, 196 illus, \$10.50 (1968)

Addison-Wesley
PUBLISHING COMPANY, INC.
Reading, Massachusetts 01867



THE SIGN OF
EXCELLENCE

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers



For years, the Institute for Scientific Information (ISI®) has been providing scientists with the information they need—in the form they need it. Now, you and your colleagues in the behavioral, social and management sciences can take advantage of ISI's vast data bank of over 4,000 domestic and foreign journals to keep you up-to-date on what's being published in your field.

Let your next journal be *Current Contents®—Behavioral, Social & Management Sciences*

For your general need to be aware of what's going on in your field, ISI publishes *Current Contents—Behavioral, Social & Management Sciences*. Subscribers to this pocket-sized weekly receive, often in advance of publication, reproductions of table-of-contents pages of over 900 domestic and foreign journals in the behavioral, social and management sciences. And an exclusive author-address directory.

Another invaluable ISI® service: *Automatic Subject Citation Alert*

If your areas of concern can be more specifically defined, ISI's Automatic Subject Citation Alert (ASCA®) will alert you, on a weekly basis, to any materials published on topics which fall within your personally-programmed interest profile areas.

And get as a bonus *Original Article Tear Sheet Service*

ISI's Original Article Tear Sheet Service (OATS®) can provide subscribers to either of the above two services any article desired on a 24-hour turn-around basis.

Keeping up with all those journals you don't have time for is easy when you subscribe to one or more of the comprehensive services ISI has designed for busy social or behavioral scientists. Write for details today.

© 1970 ISI

ISI® INSTITUTE FOR SCIENTIFIC INFORMATION

325 CHESTNUT ST., PHILADELPHIA, PENNSYLVANIA 19106, USA
132 HIGH ST., UXBRIDGE, MIDDLESEX, UK
OTHER OFFICES IN WASHINGTON, PARIS, TOKYO
TELEX: 84-5305 CABLE: SCINFO AER-4

Please send me information on:

- ☐ *Current Contents®—Behavioral, Social & Management Sciences*
☐ Automatic Subject Citation Alert (ASCA®) ☐ Other ISI Information Services
☐ Original Article Tear Sheet Service (OATS®)

Name/Title _____

Organization _____

Address _____

City _____ State _____ Zip _____

Country _____

Please mention *THE AMERICAN ECONOMIC REVIEW* When Writing to Advertisers

Now available . . . an important new introduction to economics . . .

INTRODUCTION TO ECONOMICS

CARL BREHM, *Kenyon College*

Random House, 1970; 224 pages; \$7.50

This text presents a brief and very clear development of micro and macroeconomic theory as well as nontechnical discussion of the methods that economists use to test their theories.
INSTRUCTOR'S MANUAL

A new introduction to money and banking . . .

MONEY, CREDIT AND BANKING

WILLIAM E. DUNKMAN, *University of Rochester*

Random House, 1970; 496 pages; \$9.95

A carefully organized exposition of the influence of monetary and fiscal policies on economic growth and stability. Special attention is given to American problems and policies since 1960.

And a forthcoming reader . . .

MONEY AND BANKING

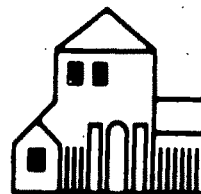
Theory, Analysis, and Policy: A Textbook of Readings

Edited by SID MITTRA, *Oakland University*

Random House, June 1970; 704 pages; \$5.95

Arranged to demonstrate that monetary theory and policy is evolutionary in nature, these readings present the whole range of thinking on important topics.

MIERNYK is coming.



RANDOM HOUSE

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

THE RANDOM HOUSE PRIMER SERIES IN ECONOMICS

paperbound editions under the general editorship of PETER L. BERNSTEIN, The New School for Social Research and Bernstein-Macauley, Inc.

The latest titles in this series are:

A PRIMER ON UNITED STATES ECONOMIC HISTORY

HARVEY C. BUNKE, *Indiana University* • 1969; 192 pages; \$2.25

A PRIMER ON PRODUCTIVITY

SOLOMON FABRICANT, *New York University* • 1969; 224 pages; \$2.25

A PRIMER ON EMPLOYMENT AND WAGES

Second Edition

WALTER GALENSON, *Cornell University* • 1969; 160 pages; \$2.25

A PRIMER ON MONOPOLY AND COMPETITION

WILLARD F. MUELLER, *University of Maryland* • 1969; 224 pages; \$2.25

Just out!

A PRIMER ON BUSINESS FORECASTING

With a Guide to Sources of Business Data

LEONARD SILK, *The Brookings Institute, and*

LOUISE M. CURLEY, *Scudder, Stevens & Clark, New York* • 1970; 164 pages; \$2.50

THE ECONOMIC SYSTEM

ROGER W. WEISS, *University of Chicago*

Random House, 1969; 128 pages; \$2.50 paperbound

ELEMENTARY LINEAR PROGRAMMING

CHARLES DAVID THROSBY, *Macquarie University*

Random House, 1970; 256 pages; \$9.95

QUANTITATIVE RESEARCH METHODS FOR BUSINESS AND ECONOMICS

HOWARD L. BALSLEY, *Texas Tech University*

Random House, 1970; 352 pages; \$9.95

The College Department, 201 East 50th Street, New York, New York 10022

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

READ THE FUTURE OF AMERICAN BUSINESS

EDUCATION AND JOBS: The Great Training Robbery,
by Ivar Berg

A noted sociologist uses data from many sources to explode the myth that higher education is essential to success in employment.

\$7.50

BEYOND AUTOMATION:
Managerial Problems of an Exploding Technology (Reissue),
by John Diebold

A challenging analysis on the impact of automation on our economy, particularly on business management, with emphasis on the problems and opportunities created by an accelerating technology.

\$8.50

BUSINESS AND SOCIAL PROGRESS:
Views of Two Generations of Executives,
edited by Clarence C. Walton

Based on a symposium sponsored by the Committee for Economic Development, this book explores the involvement of American business in social issues.

\$5.95

REPORT FROM WASTELAND:
America's Military-Industrial Complex, by Senator William Proxmire

In this explosive book, the Senator from Wisconsin documents the ways men in and out of government are mismanaging the vast sums of money appropriated for military programs.

\$6.95

THE BUREAU OF THE BUDGET
by Percival Flack Brundage

How the richest nation in the world balances its budget is the theme of this new book about one of the least known and least understood Federal agencies.

\$10.00

BUSINESS DECISIONS AND TECHNOLOGICAL CHANGE
by John Diebold

This thought-provoking book examines the effects of technology on corporate structure and competition, and provides guidelines for managing technology.

\$10.00



PRAEGER PUBLISHERS

111 Fourth Avenue, New York, N. Y. 10003

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

new 1970 titles

from SOUTH-WESTERN



ECONOMICS

Hallstones, Xavier University
Brennan, Brown University

... a book of depth and contemporary concern featuring careful, lucid writing by competent economists, attractive graphics to assist the student in visualizing complex relationships, and a total teaching/learning system to insure maximum student involvement.



AN ECONOMIC HISTORY OF THE UNITED STATES

Tuttle and Perry, University of Florida

... a different and, in some ways, unique approach to American economic history providing an analysis in economic terms of many problems not covered by traditional economic historians and an integrated view of economic growth based on national income analysis and on the chronological-institutional approach.

New 1970 Revisions, too ...



THE EVOLUTION OF ECONOMIC THOUGHT, 2nd Edition

Kuhn, University of Nebraska

... a revision of a scholarly textbook designed to acquaint the student with the leading economic theorists who have influenced economics as a body of scientifically developed propositions.



CONTEMPORARY PROBLEMS AND ISSUES, 2nd Edition

Hallstones, Xavier University
Martin, Canisius College
Mastrianna, Xavier University

... a revision of a popular softbound textbook designed to help the student make the transition from the study of theoretical material to a thoughtful analysis of the issues besetting our economic system.

The Mark of a



Good Textbook

SOUTH-WESTERN PUBLISHING CO.

(Specialists in Business and Economic Education)

Cincinnati, OH 45227 • Chicago, IL 60644 • Dallas, TX 75202

New Rochelle, NY 10802 • Burlingame, CA 94010

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

from Brookings

The Political Economy of Prosperity by Arthur M. Okun

A former chairman of the President's Council of Economic Advisers reappraises the key issues of overall economic performance and policy: sustaining prosperity, combining it with price stability, and channelling the dividends of growth into high-priority uses. He examines the stabilization policies of the sixties, explains the influence and impact of the "new economics," focuses on the sources of error and imperfection in policy decisions, and points toward an evolving strategy of stabilization policy for the future. The influence of economists on public policy is analyzed and evaluated.

1970 152pp \$4.95

The Politics and Economics of Public Spending by Charles L. Schultze

"A really important book rarely appears. But the publication . . . of Charles Schultze's 1968 H. Rowan Gaither Lectures is such an event."

The Times (London)

The author examines the relation between the analytical and political approaches to budgetary decisions. After reviewing the evolution of the federal budget, he discusses the planning, programming, and budgeting system (PPB), seeks to define its role in political decision-making, and suggests specific uses for systematic analysis within the present-day political and institutional framework.

1969 143pp paper \$1.95 cloth \$4.95

Financing State and Local Governments Revised Edition

by James A. Maxwell

The fiscal problems created by the rapid growth of state and local spending for education, roads, welfare, and other public services are analyzed in this widely adopted book. The author also shows why the present federal-state-local balance of power will be threatened unless new revenue sources or expanded tax bases are found. The diverse ways in which states and localities raise their money, and their problems and opportunities, are examined in historical perspective.

1969 275pp paper \$2.95 cloth \$8.75

Federal Budget Policy, Revised Edition

by David J. Ott and Attiat F. Ott

Now available in a new edition, this book describes and analyzes the history of federal expenditure and tax policy, the budget process, and budget concepts. Special features of the revised edition are a thorough description of the new form of unified budget recommended by the President's Commission on Budget Concepts in 1967 and adopted for the presentation to Congress of the budget for fiscal year 1969, and a report on the impact of the planning-programming-budgeting system (PPB).

1969 182pp paper \$2.25 cloth \$5.00

Federal Tax Policy

by Joseph A. Pechman

"... one of the most important works on taxation that has come out so far this decade." *Business Week* "... a gem as an introduction to the subject—a shining combination of good writing and good thinking". *National Tax Journal*

1966 321pp paper \$2.45 cloth \$5.00

**Traders and Diplomats:
An Analysis of the Kennedy Round of
Negotiations under the General
Agreement on Tariffs and Trade**
by Ernest H. Preeg

A member of the U.S. negotiating team at Geneva discusses the origins of the Kennedy Round and examines the negotiations chronologically. He analyzes the results as they pertain to tariff cuts in the industrial and agricultural sectors, to the reduction of tariff dispersion and tariff disparities, to developing countries, and to food aid commitments. He suggests some of the implications for future trade policy.

1970 320pp \$8.75

**The Dilemma of
Freight Transport Regulation**
by Ann F. Friedlaender

Increasingly, the goals and methods of public regulation of the freight transportation industry are being called into question, with the need for change widely accepted. This book, the first in the Brookings series of Studies in the Regulation of Economic Activity, reviews the existing regulatory policies, their origins and deficiencies and the probable effects of a number of alternative policies.

1969 218pp \$8.75

The International Airline Industry
by Mahlon R. Straszheim

This ninth book in the Brookings Transport Research Program series analyzes the international airlines, discusses the problems which impair efficiency, and suggests policies which would improve the operation of the industry and promote the best interests of the traveling public.

1969 297pp \$7.50

**Economic Evaluation of Urban Renewal:
Conceptual Foundation of
Benefit-Cost Analysis**
by Jerome Rothenberg

"... probably the best available a priori analysis of urban renewal and its possible consequences. It clearly deserves careful reading..." *The Journal of Business* "This is a top-notch example of cost-benefit analysis..." *Management Science*

1967 277pp paper \$2.50 cloth \$6.75

Poverty and Discrimination
by Lester C. Thurow

An investigation of the differing causes of poverty among blacks and whites with particular emphasis on the ways in which racial discrimination affects the efficacy of various antipoverty measures.

1969 214pp \$8.75

Education and Poverty
by Thomas I. Ribich

A major assumption of the war on poverty has been that education and training are especially effective means of helping people to lift themselves out of poverty. By applying cost-benefit analysis to several educational programs undertaken as anti-poverty measures, the author tests that assumption and raises doubts about the scope of its validity.

1968 163pp \$5.00

Negative Taxes and the Poverty Problem
by Christopher Green

"... an excellent appraisal of the whole negative taxation question..." *Monthly Labor Review*

1967 210pp paper \$2.50 cloth \$8.75

The Brookings Institution

1775 MASSACHUSETTS AVENUE N.W. WASHINGTON D.C. 20036

Investment Interest and Capital

J. Hirshleifer, University of California at Los Angeles

Offering a unified treatment of the theory of intertemporal decision and equilibrium, this new book shows the natural extension and generalization of standard price theory into the domain of time. This generalization also extends to uncertainty theory. The exposition is elementary, correlating the insights derived from calculus with geometrical and traditional analysis.

The book incorporates a variety of techniques: calculus, geometry and literary analysis. Numerous examples reinforce the text. Results and insights derived from dif-

ferent techniques are compared and correlated. Part I covers intertemporal resource allocation under certainty; Part II broaches the more difficult topic of decision and equilibrium under uncertainty. The author clearly spells out the crucial role played by market imperfections, and shows why this is a vital area for future research. Emphasis is on the "real" theory of interest.

February 1970 336 pp. \$9.95 (50295-5)
For more information write Box 903,
Prentice-Hall, Englewood Cliffs, N.J. 07632

THE JOURNAL OF ECONOMIC EDUCATION

**the only publication devoted
exclusively to reporting all major
developments in teaching economics.**

- Original articles extending understanding of the teaching process.
- Summaries of significant contributions appearing elsewhere.
- Special departments: research in progress . . . book reviews . . . announcements of conferences, seminars and fellowship programs . . . notices of personnel available in economic education.

Among contributors: G. L. Bach, *Stanford University*, Kenneth Boulding, *University of Colorado*, Keith Lumsden, *Stanford University*, Campbell McConnell, *University of Nebraska*, Donald Paden, *University of Illinois*, George J. Stigler, *University of Chicago*, R. Sztreter, *University of Birmingham (England)*, Rudolph Trenton, *Oklahoma State University*.

The Journal of Economic Education is published in April and November by the Joint Council on Economic Education with the advice of members of the American Economic Association's Committee on Economic Education. Annual subscription, \$3; individual issues, \$2; 10 or more subscriptions to one address, \$2.40 per subscription. Foreign subscriptions add \$1.50 per year. All subscriptions must be prepaid.

Send orders, accompanied by check made payable to *The Journal of Economic Education*, to:

1212 AVENUE OF THE AMERICAS / NEW YORK, NEW YORK 10036

THE JOURNAL OF ECONOMIC EDUCATION

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

**Joseph
Dorfman**

THE ECONOMIC MIND IN AMERICAN CIVILIZATION 1606-1933

5 vols. (1946-1959)

This monumental and comprehensive work—the product of a quarter century of research, study, and reflection—is of unparalleled significance in the history of American economic thought. Moreover, in its broad analytic perspective of the social, political, and economic background against which the development of economic ideas took shape, *The Economic Mind in American Civilization* remains a fundamental study in American culture.

In his detailed exposition of the history of the American contribution to economic thinking, Professor Dorfman, drawing from unpublished as well as published sources, has given attention to the contributions not only of the formal thinkers and recognized expositors of the economic discipline, but also of informal commentators, public figures, non-academic and popular influences.

"His work has been remarkably thorough—
superseding that of his most notable
predecessor, Vernon L. Parrington, in range
and objectivity."

Wesley C. Mitchell

"Professor Dorfman is a prodigious scholar..."
John Kenneth Galbraith

"Incomparably the best thing ever written on
the development of economic thought and
doctrine...extraordinary readability."

Alvin Johnson

"The volumes blaze a new trail. They make
available a vast amount of material pertinent
to an understanding of American civilization...
Veblen would have liked this book."

**Ralph Henry Gabriel,
The New York Times Book Review**

"An important addition to the shelves of
economic literature, and comforting proof that
economists have made a lot of progress in
the last couple of centuries."

The New Yorker

CONTENTS

Volume I

Book One: Colonial America

Book Two: From Independence to Jackson
Part One: The Emergence of Traditions of
"Free Enterprise."

Volume II

Part Two: The Emergence of Systematic Economics
Book Three: From Jackson to the Civil War

Volume III

Part I: The Era from 1865 to 1880
Part II: The Expansion of Economic Thought
Part III: The Heartbreaking Nineties
Part IV: The Promise of the New Century

Volume IV

Part I: The Era of Adjustment and Progress, 1918-1929

Volume V

Part I: The Era of Adjustment and Progress, 1918-1929
(continued)
Part II: The Great Depression, 1929-1933

Now available for the first time in many years as a complete set!

Vol.'s I & II, 1606-1865.....	\$20.00
Vol. III, 1865-1918.....	12.50
Vol.'s IV & V, 1918-1933.....	20.00
For the set of five volumes.....	\$47.50



AUGUSTUS M. KELLEY, Publishers
1140 Broadway, New York, N. Y. 10001

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

harper torchbooks

New and forthcoming

CHRIS ARGYRIS

Personality and Organization *The Conflict Between System and the Individual*

"[This] integrated review will produce, I hope, a few basic foundations for . . . a systematic framework for the field of organizational behavior."—CHRIS ARGYRIS, in the Introduction

TB/1495 \$1.95

DAVID BRODY

Steelworkers in America *The Nonunion Era*

"A lively, well-written, and carefully documented account of . . . labor in steel in the three decades that preceded the great depression."—JOEL SZYMAN, *Journal of Political Economy*

TB/1485 \$2.95

W. ARTHUR LEWIS

Theory of Economic Growth

"A landmark in the contemporary approach to economics."—*The Observer* (London). "An extremely important book. . . A working handbook for statesmen as well as a detached study."—*The Economist*

TB/1492 \$2.45

WILLIAM F. WHYTE

Money and Motivation *An Analysis of Incentives in Industry*

"Whyte has contributed considerable insight to our understanding of incentives and worker motivation."—IRVIN SOBEL, *The American Economic Review*. (Researches in the Social, Cultural, and Behavioral Sciences, edited by Benjamin Nelson.)

TB/1494 \$1.95



1817

Complete catalog of Harper Torchbooks [733 volumes now in print] available on request from Dept. 51, HARPER & ROW, Publishers, Inc., 49 E. 33d St., New York, N.Y. 10016

CONTRIBUTIONS TO ECONOMIC ANALYSIS

Editors: J. TINBERGEN, D. W. JORGENSON, J. WELBROECK and J. JOHNSTON

New Volumes:

Vol. 60: A Structural Model of the U.S. Balance of Payments.

By M. F. J. PRACHOWNY, Queen's University, Kingston, Ontario, Canada.

1969. 176 pages. Hfl. 34.— (79 s., \$9.50)
ISBN 7204 3157 3

Vol. 61: A Contribution to the General Equilibrium Theory of Preferential Trading.

By M. C. KEMP, University of New South Wales, Kensington, Australia.

1969. 149 pages. Hfl. 28.50 (67 s., \$8.00)
ISBN 7204 3162 X

Vol. 62: Industry Production Behavior: The Order-Stock Distinction.

By D. A. BELSLEY, Department of Economics, Boston College.

1969. 250 pages. Hfl. 45.— (105 s., \$12.60)
ISBN 7204 3159 X

Further volumes are in active preparation.

A complete list of titles published in the series "Contributions to Economic Analysis" is available on request.

Vol. 63: Optimal Economic Growth.

Shifting finite versus infinite time horizon.
By M. INAGAKI, Department of Economics,
University of Montreal

1970. Ca. 200 pages. In preparation.
ISBN 7204 3161 1

Vol. 64: American and British Regional Export Determinants.

By A. L. GINSBURG, Department of Economics,
City University of New York

1969. 199 pages. Hfl. 35.— (82 s., \$9.75)
ISBN 7204 3158 1

Vol. 65: Estimating the Parameters of the Markov Probability Model from Aggregate Time Series Data.

By T. C. LEE, University of Connecticut, G. G. JUDGE,
University of Illinois, and A. ZELLNER, University of
Chicago.

1970. 254 pages. Hfl. 43.— (100 s., \$12.00)
ISBN 7204 3163 8

AMSTERDAM
P.O. BOX 3489

NORTH



HOLLAND

PUBLISHING CO.
NETHERLANDS

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

WE'D LIKE YOU TO MEET THE ARKANSAS TRAVELER.

If you happen to teach in Arkansas, that is. If not, how about meeting the Colorado Traveler, the Michigan Traveler, the Pennsylvania Traveler, etc.? They'll tell you about economics texts that say it straight. About economics texts that excite the mind. And about economics texts that do enlighten today's econ student. Economics texts from HRW. Of course.

INTRODUCTION TO MACROECONOMIC THEORY

Edward Ames, SUNY, Stony Brook
1968 / 144 pages / \$2.95 paper
(Principles of Economics Series)

INTRODUCTION TO INPUT-OUTPUT ECONOMICS

Chiou-shuang Yan, Drexel Institute of Technology

1969 / 144 pages / \$2.95 paper
(Principles of Economics Series)

MACROECONOMIC ADJUSTMENTS

John A. Carlson, Purdue University
January 1970 / 160 pages / \$2.50 paper
(Principles of Economics Series)

MODERN POLITICAL ARITHMETIC

Bruce F. Davis and Bruce F. Duncombe,
both at Georgetown University
February 1970 / 128 pages / \$3.95 paper
(tent.) (Economic Series)

MACRO: A Game of Growth and Policy

Peter H. Lindert, University of Wisconsin,
Madison

February 1970 / 72 pages / \$2.95 paper
Referee's Manual (Economics Series)

READINGS IN THE HISTORY OF ECONOMIC THEORY

Ingrid Rima, Temple University
January 1970 / 256 pages / \$5.95 paper
(tent.) (Economics Series)

A RADICAL'S GUIDE TO ECONOMIC REALITY

Angus Black
April 1970 / 96 pages / \$2.95 paper (tent.)

DECISION-MAKING AND THE THEORY OF THE FIRM

Ira Horowitz, Indiana University, Bloomington
January 1970 / 416 pages / \$10.00 (tent.)

READINGS IN THE ECONOMICS OF INDUSTRIAL ORGANIZATION

Edited by Douglas Needham, SUNY, Brockport
January 1970 / 320 pages / \$5.95 paper

THE EMERGENCE OF MANPOWER POLICY

Garth L. Mangum, University of Utah
January 1970 / 192 pages / \$3.50 paper
(Business and Society Series)

PORTFOLIO BEHAVIOR OF FINANCIAL INSTITUTIONS

William L. Silber, New York University
January 1970 / 160 pages / \$6.95
(Finance Series)

STATISTICAL ANALYSIS: With Business and Economic Applications

Ya-lun Chou, St. John's University
1969 / 816 pages / \$12.95

Instructor's Manual

Study Guide
Donald R. Plane, University of Colorado
1969 / 288 pages / \$3.95 paper

MATH AND STATISTICS FOR ECONOMISTS, Second Edition

Gerhard Tintner, University of Southern California, and Charles B. Millham,
Washington State University
January 1970 / 480 pages / \$10.95 (tent.)

For more details about these and other fine economics texts—or for a meeting with the Arkansas Traveler or any other Holt Traveler, please write to Mrs. Rhoda Villanueva, copywriter, College Promotion.



Holt, Rinehart and Winston, Inc.

383 Madison Avenue, New York 10017

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Announcement of the 1970 session of the

ECONOMICS INSTITUTE

Sponsored by the

American Economic Association

in cooperation with the

Institute of International Education and other Foreign Student Sponsoring Agencies

A SPECIAL TRANSITIONAL TRAINING PROGRAM FOR SELECTED FOREIGN STUDENTS WHO HAVE BEEN ADMITTED TO REGULAR GRADUATE STUDY IN ECONOMICS OR AGRICULTURAL ECONOMICS AT UNIVERSITIES AND COLLEGES IN THE UNITED STATES. THE PURPOSE IS TO HELP STUDENTS STRENGTHEN THEIR PREPARATION AND BECOME FAMILIAR WITH UNITED STATES ACADEMIC PROCEDURES SO THEY MAY CARRY ON THEIR GRADUATE STUDIES MORE PRODUCTIVELY AND ADVANTAGEOUSLY. THE FACULTY OF THE INSTITUTE IS RECRUITED NATIONALLY.

• THE PROGRAM INCLUDES INTENSIVE COURSE WORK—OFFERED AT SEVERAL LEVELS—IN ECONOMIC THEORY, QUANTITATIVE TECHNIQUES AND ENGLISH. THIS COURSE WORK IS SUPPLEMENTED BY SPECIAL LECTURES AND SEMINAR SERIES AND BY FIELD TRIPS AND SOCIAL AND RECREATIONAL ACTIVITIES DESIGNED TO ACQUAINT THE PARTICIPANTS WITH THE UNITED STATES ECONOMY AND CAMPUS AND COMMUNITY LIFE. INDIVIDUAL PROGRAM ASSIGNMENTS ARE DETERMINED ON THE BASIS OF A SERIES OF PLACEMENT TESTS AND INTERVIEWS.

Location and Dates of Institute

UNIVERSITY OF COLORADO
June 17 - August 19, 1970

Admission and Fellowship Information

Final admission to the Institute will be granted only to those students who can provide proof of admission to graduate work in economics or agricultural economics at a university or college in the United States for the academic year following the Institute. Preference is given to applicants who have not previously studied economics in the United States. The Institute is designed primarily for full session participants, but provision is made for the admission of a limited number of applicants for the last five weeks of the program only (i.e., July 15-August 19). Preference is given to single students and to married students who will not be accompanied by any dependents since only these will be accommodated in the Institute's residential facilities. The admission of others is not precluded but will be conditional upon the applicant being able to arrange satisfactory living facilities for himself and his family independent of the Economics Institute and prior to its opening date. Applicants and/or their academic year sponsors normally will be expected to pay the full costs of their participation in the program. However, a number of partial cost fellowships will be allotted on a competitive basis to applicants without alternative sources of full financial support. The total enrollment is limited to approximately 100 students. Further details together with admission and fellowship application forms can be obtained from the office of the director.

A.E.A. Policy and Advisory Board

HENRY ROSEVEY, Harvard University, Chairman
DOUGLASS C. NORTH, University of Washington
GUSTAV RANTIS, Yale University
VERNON W. RUITAN, University of Minnesota
JOSEPH J. SPENGLER, Duke University
JEFFREY WILLIAMSON, University of Wisconsin



Director

WYN F. OWEN, University of Colorado
Office of the Director
Economics Institute
University of Colorado
Boulder, Colorado
USA 80302
Telephone (303) 443-2211
Ext. 7337
Cable ECONINIST

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

New MARKHAM Books in Economics

NOTES AND PROBLEMS IN MICROECONOMIC THEORY

by Samuel Bowles and David Kendrick

in collaboration with Lance Taylor and Marc Roberts

This book consists of an annotated list of fundamental reading sources and a set of problems which present the basic concepts of microeconomic theory.

The problems presented here are intended to help the student become thoroughly versed in the fundamentals of the subject, while at the same time discovering for himself the strengths and weaknesses of the current state of development of microeconomic theory.

"There has been a strong shift toward the use of problem-solving as a method of instruction in the teaching of economic theory, especially at the graduate level. The Bowles-Kendrick is the first comprehensive collection of problems that can be used for instructional purposes."

Dale W. Jorgenson Harvard University

Publication May 1970

275 pages

\$3.50 List, Paper Cover

PUBLIC EXPENDITURES AND POLICY ANALYSIS

Edited by Robert Haveman and Julius Margolis

The papers in this volume form a comprehensive study of public expenditure economics. Both theory and current practice in public expenditure analysis are covered. Major sections of the book deal with conceptual problems in application of economic analysis and institutional factors affecting implementation of public expenditure decisions.

Publication May 1970

595 pages

\$9.50 List, Hard Cover / \$6.50 List, Paper Cover

REBELLION AND AUTHORITY: An Analytic Essay on Insurgent Conflicts

by Nathan Leites and Charles Wolf, Jr.

The broad aim of this book is to develop a method of analyzing insurgent conflicts which approaches the rigor achieved in analysis of strategic nuclear conflicts. Various concepts are developed, such as the "demand" for and "supply" of rebellion; the "inputs" and outputs associated with rebellion; "damage-limiting" and "profit maximizing" influences on population behavior; and the principles of target-selection as they affect the contesting sides. The concepts are then applied to specific cases of insurgent conflict both foreign and domestic.

Published February 1970

174 pages

\$5.95 List, Hard Cover

COST AND CHOICE: An Inquiry in Economic Theory

by James M. Buchanan

This new book presents a critical examination of the fundamental elements of traditional cost theory and challenges the prevailing orthodoxy that the opportunity cost which influences choice can be objectively measured. In a literate, straightforward, non-mathematical exposition, the author presents ideas that will stimulate serious reconsideration of the basic concepts of cost and choice.

Published November 1969

122 pages

\$4.95 List, Hard Cover

MARKHAM publishing company

3322 West Peterson Avenue • Chicago, Illinois 60645 • (312) 539-0626

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Just Published by the IMF

THE INTERNATIONAL MONETARY FUND, 1945-1965

Twenty Years of International Monetary Cooperation

Vol. I Chronicle, 663 pp.

Vol. II Analysis, 621 pp.

Vol. III Documents, 549 pp.

This history of the first twenty years of the International Monetary Fund, written by past and present members of its staff, attempts to recount objectively what were its origins and what have been its plans and activities.

Volume I contains a chronological account of the evolution of the Fund from the original plans formulated in 1941 to the end of 1965, showing the development of its policies and the influence on the Fund of various national viewpoints. Volume II analyzes the Fund's policies in three major fields and reviews its constitutional development. Volume III reproduces the major documents referred to in the other two volumes including some not hitherto published—notably early versions of the plans devised by Lord Keynes and Mr. Harry Dexter White.

Price: \$12.50 the set of three volumes (\$5.00 a volume if sold separately), or the equivalent in most other currencies.

Address inquiries to

Sales Manager

International Monetary Fund

19th and H Streets, N.W.

Washington, D.C., 20431 U.S.A.

New

The Political Economy of Change

Warren F. Ilchman and Norman Thomas Uphoff

In this attempt to provide a social science superior to present political science and more relevant than economics for the task of analyzing public policy, especially in relation to developing countries, the authors explore the functions and characteristics of political resources, political exchange, and political productivity. This perspective permits them to compare the efficiency of various public policies and to estimate their costs. They believe that they can answer "How much is enough?" when the decision to invest in political and administrative infra-structure is made, and can incorporate the requirements of political optimality and feasibility into the same system that considers economic optimality and feasibility. For the social scientist in search of a scheme that can deal with any regime, this volume should prove stimulating and refreshing.

1969 LC: 71-81743 384 pages \$8.50



from California

University of California Press • Berkeley 94720

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Recent and forthcoming—

Essays in European Economic History

F. Crouzet, W. H. Chaloner, and W. Stern, editors

This volume makes available previously untranslated or difficult to obtain articles on European economic history, covering a period from the French Revolution to the First World War, a period that saw the development of European industrialization. The book includes essays from several countries in Western and Central Europe, essays chosen for their quality, that offered new theories when they appeared or revised traditionally accepted ideas and that are not purely factual but deal with problems of general interest. The volume thus offers an eclectic selection of viewpoints on nineteenth and early twentieth century economic developments in continental Europe.

January, 1970 272 pages \$12.00

and revised editions of major works by Paul Einzig—

A Textbook of Foreign Exchange

A discussion of exchange control, the Euro-currency markets, the recently developed interbank sterling market, foreign exchange theories and policies, and the history of foreign exchange. The new edition includes a detailed study of the changes brought about by the devaluation of sterling in 1967.

Second edition, 1969 272 pages \$7.50

The Euro-Bond Market

The tremendous expansion of the European market in foreign bonds since the publication of the first edition of this book has led to this revised and enlarged edition. The market in foreign dollar loans in Europe described in the original edition has come to be known as the Euro-Bond market, hence the change in title from *Foreign Dollar Loans in Europe*. New chapters cover the emergence of a secondary market, the increase in D. mark bond issues, the changing role of the United States from a lender in Europe to a borrower from Europe, and the growing popularity of convertible bonds.

January, 1970 224 pages \$9.00

The Euro-Dollar System

Although transactions in Euro-dollars and in other foreign currency deposits have assumed considerable importance during recent years, this book is the sole attempt to describe the practice fully and to analyze its theoretical implications. This fourth edition offers updated material and a new chapter describing the spectacular increase in the volume of American borrowing in Europe and its effect on interest rates.

May, 1970 232 pages \$12.00

Foreign Exchange Crises

A detailed analysis of economic crises in general, discussing the causes of such crises, their symptoms, their effects, and the means governments employ to deal with them. This new edition considers major events that have occurred since the book first appeared, including the devaluation of sterling and of the franc, the adoption of a two-tier gold system, and the strengthening of the dollar.

May, 1970 240 pages \$7.50

ST. MARTIN'S PRESS

College Department

175 Fifth Avenue

New York, New York 10010

The Great Richmond Terminal

A Study in Businessmen and Business Strategy

By MAURY KLEIN, *University of Rhode Island*. xii, 323pp., maps, tables, appendixes, bibliog., index. \$9.50
(Eleutherian Mills-Hagley Foundation)

This case history of the Richmond Terminal, one of America's first pure holding companies, documents the process of economic development in the South during the post-Civil War era. The railroad was the first large southern industry to arise after the war and the first to confront the major decision-making problems of organization, administration, and mobilization of capital. The Richmond Terminal is of particular significance because its history corresponds to the basic pattern of the railroad's development and it offers insight into the evolution of corporate and financial institutions. Associated with the history and growth of all the towns throughout the south, the Richmond Terminal is of further interest because, in its existence today as the Southern Railway Company, it remains the dominant system in the South.

Legitimacy of the Business Corporation

By JAMES WILLARD HURST, *University of Wisconsin*. 181 pp., bibliog., index. \$6.50

These three essays deal with public policy and the business corporation in the United States from 1780 into the 1960's. The study follows the almost two hundred years of public policy which has concentrated on making the corporation a "legitimate instrument of business energy and ambition." Full documentation is included with detailed references to all relevant legal materials. This examination of the legitimacy of privately organized power is the most complete study of this important force available.

University Press of Virginia

Charlottesville

FISHERY ECONOMISTS are required by FOOD AND AGRICULTURE ORGANIZATION of the UNITED NATIONS in various countries

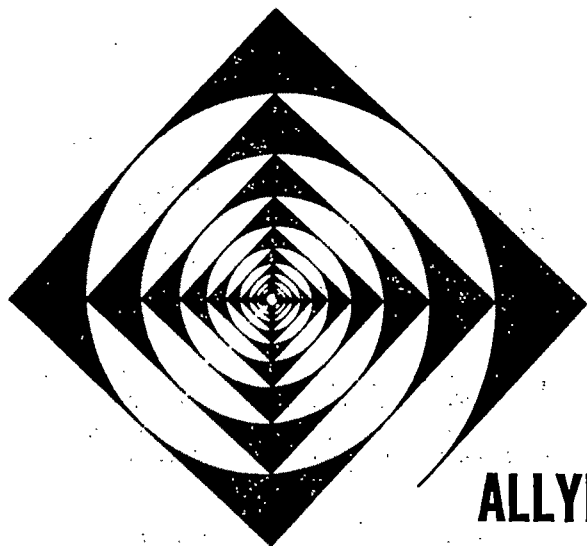
- IF YOU HAVE:
- Degree in economics or mathematics and economics from a university of recognized standing
 - Five to seven years experience, depending on the assignment, as an economist in government departments, private enterprise or in international organizations, especially in the field of fisheries
 - A very good knowledge of one of the three official languages of FAO (English, French and Spanish)
 - Ability to establish and maintain good working relations with people of different social and cultural backgrounds

WE CAN OFFER: Annual Salary ranging from US \$9,666 to 15,119 net (tax free), plus post Cost of Living Adjustment, other allowances and liberal fringe benefits

IF INTERESTED: Send résumé to:
Chief, Recruitment Office
FAO of the United Nations
Viale delle Terme di Caracalla
00100 Roma, Italy

Quoting Reference HF/FI/3

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers



NEW from
ALLYN AND BACON, INC.

Ascent to Affluence: A History of American Economic Development. Charles H. Hession and Hyman Sardy, both of Brooklyn College, 1969, 6 x 9, 912 pp. This unique text integrates the theories of economic growth with detailed historical analysis of the entire American economic experience from colonial times to the present. Each of the ten parts of the book is preceded by an Introduction that points out one or more theories of growth as a possible analytical framework for the understanding of the period in question. Succeeding chapters describe the main sectors of growth in the period, analyze the applicability of the theory of the historical data, and interpret the political and cultural processes which accompanied the economic changes.

NEW • 1970/NOW AVAILABLE. Mathematical Models in Microeconomics. Coldwell Daniel, III, University of Houston, 1970, 6 x 9, 228 pp. This new book presents a clear, rigorous, and well-organized development of the basic models of microeconomic analysis. Precise exposition helps the reader to understand microeconomics step-by-step; the use of mathematical logic encourages consistent applications of fundamental economic theory.

NEW—1970. Quantitative International Economics. Edward E. Leamer and Robert M. Stern of the University of Michigan. 1970, 6 x 9, est. 224 pp. The object of this book is to set forth on a reasonably advanced level the various methods for quantitative measurement of the most important relationships at issue in the areas of the balance of payments and International trade and welfare, to give some of the recent research in the area, and to indicate directions for new research.

ALLYN AND BACON, INC. 470 Atlantic Avenue
Boston, Massachusetts 02210

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

the new economics books

THE WELFARE STATE: U. S. A. An Exploration In and Beyond the New Economics

Melville J. Ulmer, University of Maryland
203 pages, Paper, March 1970
(Cloth, 1969, \$4.95)

- describes the welfare state structure and its theoretical foundations, including the new economics
- analyzes our pressing economic problems and the government's efforts to solve them
- examines the Keynesian principles of economic planning in operation, and
- develops a new plan for treating unemployment and poverty and providing an unprecedented measure of stability for the overall economy of the nation.

An excellent supplement for introductory economics and for reading by economics students at all levels.

MACROECONOMICS: Selected Readings

**Edited by Walter L. Johnson and
David R. Kamerschen**,
both of the University of Missouri
520 pages, Paper, Late Spring 1970

- thirty-seven selections from the most recent literature show macroeconomics in action
- such topics as wage-price guidelines are included
- a strong section on policy allows application of preceding theoretical material
- editorial part Introductions give cohesion to the volume
- exercises for the student are provided
- designed and organized for the typical intermediate macro course; suitable also as background in the first graduate level course.

FINANCIAL INSTITUTIONS AND MARKETS

Murray E. Polakoff, New York University,
and Others
650 pages, Late Spring 1970

- a distinguished, contemporary text for courses in money and banking and financial institutions and markets
- twenty-eight outstanding specialist-authors contribute original chapters to a skillfully organized and well integrated volume
- particular care has been taken by each author to achieve uniformity with the overall design of the book
- the conceptual framework is based on flow-of-funds analysis as the central organizing device which best serves to integrate theory, description, empirical data, and policy
- a distinctive feature is the extensive analysis and comprehensive treatment given to the principal financial intermediaries, in a total financial-system approach
- text organization makes the book equally appropriate for one-semester or longer courses in money and banking or financial institutions and markets.

**Houghton
Mifflin**

Publisher of the American Heritage Dictionary of the English Language

Boston, Atlanta, Dallas, Geneva, Ill., New York, Palo Alto

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Announcing

THE WADSWORTH SERIES IN LABOR ECONOMICS AND INDUSTRIAL RELATIONS

edited by *Lloyd Ulman* and *George Strauss*, Institute of Industrial Relations,
University of California, Berkeley

This series was inspired by the realization that some of our most serious social issues both arise within and profoundly modify the labor market and its institutions.

Among other high interest topics, the paperback volumes in the series will treat such problems as: the rights and welfare of ethnic minorities; unionization of white-collar and professional people; economic growth and unemployment; automation and job training; antitrust, conglomerates, and coalition bargaining; union democracy and the collective bargaining process; and comparative labor movements.

Our first volume:

THE INVESTMENT IN HUMAN CAPITAL

by *Lester C. Thurow*, Massachusetts Institute of
Technology. Available in July 1970.

About 125 pages. 5% x 8%. Paperbound.

Just published

ECONOMICS IN AMERICAN SOCIETY

by *John A. Orr*, Florida Atlantic University, and *Donald T. Savage*, University of Maine

Students complain that their introductory course in economics is not "relevant." Instructors object to books that restrict course flexibility or pre-empt the major role of the instructor in the teaching process.

We have just published an introductory text that acknowledges these complaints. We are excited about the stated objectives of the book: to make economics interesting, to cause students to ponder some of the controversial economic and social questions of our age, and to interest students in developing their economic understanding by taking additional courses.

We are equally enthusiastic about the unique teaching tools employed in the text: CHAIN REACTION DIAGRAMS, allowing the instructor to move further into economic controversies than is usually possible in a brief core text; AUTHOR OPINION SECTIONS, stimulating students to formulate their own position on controversial economic problems; and QUESTIONS that encourage the student to apply learned concepts to practical problems rather than merely return printed words. February 1970. 384 pages. 7½ x 9¼. Clothbound.

For your examination copies of these books and additional information on the Wadsworth Series in Labor Economics and Industrial Relations, please write to Box AER-3,

WADSWORTH PUBLISHING COMPANY, INC.

Belmont, California 94002

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

New and Forthcoming

TIME SERIES ANALYSIS, FORECASTING AND CONTROL

by GEORGE E. P. BOX, University of Wisconsin and GWYILYM M. JENKINS, University of Lancaster. This book is concerned with discrete time series and dynamic models and their use in forecasting and control. The techniques are all illustrated with real examples using economic and industrial data. *June, 1970; 500 pages, \$18.95 (est.)*

UNCERTAINTY AND ESTIMATION IN ECONOMICS In Three Volumes

by D. G. CHAMPERNOWNE, Cambridge University. Describes the use and interpretation of imperfect knowledge and imperfect information within the social sciences, particularly economics. Volume I covers preference and probability orderings, measurements plus varied models and random samplings, and Bayesian methods of estimation. Volume II continues with the experimental method in economics, significance tests, point estimators, interval estimators, social surveys, time series, sequential sampling, regression models and econometrics, and equilibrium of frequency distributions. Volume III considers situations of complete uncertainty, choice in risky situations, portfolio selection, measurement of risk-aversion and diminishing marginal utility, project analysis and planning under uncertainty, and the effects of uncertainty in public affairs. *Volume I, January, 1970, 288 pages, \$9.75; Volume II, 428 pages, January, 1970, \$14.75, Volume III, 112 pages, January, 1970, \$5.00.*

DYNAMIC PROGRAMMING

by D. J. WHITE, University of Strathclyde, Glasgow. Designed to remove previous misconceptions of Dynamic Programming and the attendant emphasis on its heavy computational aspects, this book provides a clear understanding of the positive contributions of D.P., its advantages and disadvantages, and its practical applications. *1969, 230 pages, \$11.95*

THE ESTIMATION OF DISTRIBUTED LAGS

by PHOEBUS J. DHRYMES, University of Pennsylvania. The first integrated presentation of distributed lags as developed over the years. It deals with the formulation, motivation, and empirical applications of distributed lag models. The algebraic structure of lag operators is developed, and their usefulness in dealing with distributed lags is demonstrated. The major distributed lag models such as the geometric, polynomial and general rational models are developed and efficient estimation procedures are devised for their parameters. Methods of discriminating among the alternative distributed lag models are also developed. *1970, 400 pages, \$12.95 (est.)*

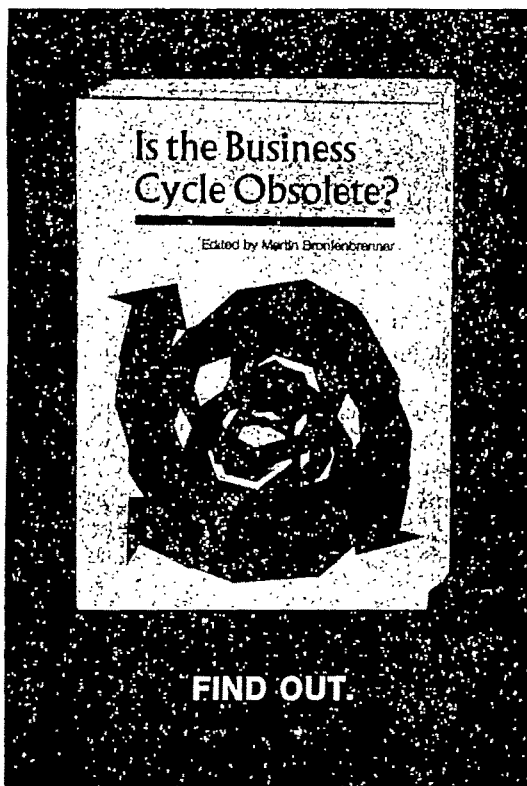
COLLECTIVE CHOICE AND SOCIAL WELFARE

by AMARTYA SEN, Delhi University. The problem of aggregating the preferences of individuals into a collective preference is the subject matter of this book. This is studied in a very general form and as such the results should be of relevance to welfare economics, theory of economic policy and planning. Integrates the relevant materials in economics, philosophy, and politics in so far as they contribute to an understanding of the problem of collective choices. *1970, 256 pages, \$15.00 (est.)*

 **HOLDEN-DAY, INC.**
500 Sansome Street

San Francisco, California 94111

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers



You'll find most of the economists represented here agreed that the business cycle as we know it is generally a theory of the past; particularly as it regards major depressions like 1873 and 1929.

These papers and discussions concern the business cycle experience of over a dozen countries in the Western and Socialist worlds and Japan. They were presented in April 1967 to the Social Science and Research Committee's conference on Economic Stability in London.

"Is the Business Cycle Obsolete?", edited by Martin Bronfenbrenner, Professor of Economics, Graduate School of Industrial Administration, Carnegie-Mellon University, has 496 pages of other interesting material, too. For instance, it contains three econometric papers about dynamic models used as devices for nipping real-world cyclical movements in the bud and as subjects of statistical research. It is also the only comprehensive English publication of the conference.

Read the chapter headings below and see what the book covers. Then send your check or money order for \$12.50 with the convenient coupon. We'll pay the postage on all paid orders.

We'll also give you 10 days to return the book for a complete refund.

That's in case you don't like what you find out.

CHAPTER HEADINGS:

- | | |
|--|---|
| I. The Experience of the Western World and Japan. | III. Economic Fluctuations under Socialism. |
| II. International Transmission of Economic Fluctuations. | IV. Econometric Studies. |
| | V. Business Cycle Policies. |
| | VI. Discussion and Debate. |

wiley!

SAVE MONEY!

Enclose your check and Wiley pays all delivery and handling charges.

TO: **WILEY-INTERSCIENCE**, Dept. 216

A Division of John Wiley & Sons, Inc., 605 Third Avenue, New York, N.Y. 10016

Gentlemen:

Please send me _____ copies of "Is the Business Cycle Obsolete?" @ \$12.50 each. I enclose \$_____ and understand that Wiley-Interscience pays all delivery and handling charges. If I am not satisfied with the book(s) I may return the order within 10 days for a full refund.

☐ Bill me.

Name _____

Address _____

City _____ State _____ Zip _____

Please mention **THE AMERICAN ECONOMIC REVIEW** When Writing to Advertisers

From Cambridge

INDUSTRIALIZATION IN AN OPEN ECONOMY

Nigeria 1945-1966

PETER KILBY

This analysis of the process of industrialization in Nigeria goes beyond questions of allocational efficiency to underlying organizational and institutional factors.

Subjects discussed include the performance of the export sector, the implications of the developing petroleum industry, the importance of foreign enterprise, the role of government, etc. In examining the development of skills and indigenous enterprise, the author questions conventional development theory.

\$14.50

READINGS IN THE CONCEPT AND MEASUREMENT OF INCOME

Edited by R. H. PARKER and G. C. HARCOURT

Brings together some of the most important writings in economic literature on the subject. An introduction by the editors provides an overall view of the concept of income.

The book is primarily for practitioners and students of economics, with several contributions on the borderline between economics and accounting.

Cloth \$12.50 Paper \$3.45

PRODUCTIVITY IN TRANSPORT

B. M. DEAKIN and T. SEWARD

An empirical study of productivity between 1952 and 1965 in six sectors of public transport and communication in the United Kingdom: railways, public road passenger transport, road haulage contracting, sea transport, port and inland water transport, air transport, and postal services and telecommunications. (*Department of Applied Economics Occasional Paper 17*)

Cloth \$7.50 Paper \$3.95

DECISION, ORDER AND TIME IN HUMAN AFFAIRS

Second Edition

G. L. S. SHACKLE

The second edition has a new preface, additions to the bibliography, and a new chapter introducing in three distinct but closely connected forms the notion of "elasticity of surprise," a measure of the sensitivity of response to disconcerting events.

\$8.50

From Cambridge

THE NEOCLASSICAL THEORY OF PRODUCTION AND DISTRIBUTION

C. E. FERGUSON

The first section of this complete and systematic study gives full coverage of microeconomic neoclassical theory. In the second part Professor Ferguson examines the "alternative" theories of Kalecki, Kaldor, Joan Robinson, Kennedy, Pasinetti and others.

The controversy over the way in which technical progress takes place is examined with regard to the recent work of Samuelson, Solow and Arrow.

\$14.50

MARSHALL, MARX AND MODERN TIMES

The Multi-Dimensional Society

CLARK KERR

Professor Kerr examines the economic views of society of Alfred Marshall and Karl Marx, and studies the multi-dimensional character of modern society, considering the varied groups within it.

He pays particular attention to the American educational system and sets forth his ideas on present-day student revolt.

"A timely book, of interest to those who seek solutions to current problems."
—*Library Journal*

\$4.95

POVERTY IN BRITAIN AND THE REFORM OF SOCIAL SECURITY

A. B. ATKINSON

This study is concerned with the adequacy of the present British social security provisions for low income families and with evaluating the cost and effectiveness of possible reforms. The proposed reforms discussed include a negative income tax, a social dividend scheme, new child allowances, more generous old age pensions, and the replacement of National Insurance contributions by a social security tax. (*Department of Applied Economics Occasional Paper 18*)

Cloth \$7.00 Paper \$3.25



CAMBRIDGE UNIVERSITY PRESS

32 East 57th Street

New York, N.Y. 10022

URBAN ECONOMICS Theory, Development, and Planning

Edited by William H. Leahy, *University of Notre Dame*,
David L. McKee, *Kent State University*, and
Robert D. Dean, *Memphis State University*

This reader is the first presentation of the best theoretical work done to date in urban economic analysis. *Urban Economics* supplies the most sophisticated and recent tools for urban analysis and includes articles which represent all the important theoretical bases currently accepted as integral parts of urban economics: location and spatial factors, urban land use, planning, central-place theory, and urban economic development.

February, 1970 352 pages (approx.)

paper, \$4.95
cloth, \$5.95

REGIONAL ECONOMICS Theory and Practice

Edited by David L. McKee, *Kent State University*,
Robert D. Dean, *Memphis State University*, and
William H. Leahy, *University of Notre Dame*

Regional Economics presents classic articles from economists of all types who have contributed to regional economic theory: economic theorists, land economists, industrial location theorists, and development economists. The book begins with an introductory section that defines regional economics and its tools. The largest portion of the book is devoted to the major problems at the regional level—expansion and integration. The last part is concerned with the making of regional economic policy.

April, 1970 304 pages (approx.)

paper, \$4.95
cloth, \$5.95

SPATIAL ECONOMIC THEORY

Edited by Robert D. Dean, *Memphis State University*,
William H. Leahy, *University of Notre Dame*, and
David L. McKee, *Kent State University*

This reader is evidence of the growing awareness among economists that a comprehensive economic theory should consider both space and time dimensions, and that the locational factors of economic activity have a significant bearing on our urban and regional economies. The book begins with an overview of spatial economics in its two major aspects—partial equilibrium and general equilibrium analysis—and investigates the subdivisions of partial equilibrium analysis—least-cost theory, locational interdependence, market-area analysis, and locational equilibrium analysis. It then focuses on general equilibrium theory, and concludes with a synthesis of the leading theories on spatial economic structure.

February, 1970 384 pages (approx.)

paper, \$4.75
cloth, \$5.95

ENTREPRENEURSHIP AND ECONOMIC DEVELOPMENT

By Peter Kilby, *Wesleyan University*

This unique volume is a carefully structured selection of essays on the role of entrepreneurship in economic development. Featuring a substantial theoretical essay by the author, this group of articles is the first to present all the major theories of entrepreneurial supply and the best applied studies carried out in underdeveloped countries.

June, 1970 448 pages (approx.)

\$11.95 text.

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

New Perspectives in Economics

Economic Progress and the Developing World

By Edward Marcus, The City University of New York

Mildred Marcus, The City University of New York

A closely reasoned, carefully written critique of the social, economic (within the nation itself and in relation to other nations and the world economy), and cultural factors which aid and inhibit economic development in the less developed nations. The book provides an excellent introduction to the theory and problems of development. No one solution to the development problem is emphasized. Rather the authors give the reader the "feel" of an underdeveloped country and the magnitude of the problems that must be overcome in order to start the development process. *Summer 1970, approx. 128 pages, illus., softbound \$2.95 (tent.)*

Urban Analysis

Readings in Housing and Urban Development

Edited by Alfred N. Page, University of Washington

Warren R. Seyfried, University of Washington

Focusing on the important economic aspects of the role of housing in urban development, this collection of scholarly and topical, theoretical and empirical, papers provides thorough coverage of the work accomplished in the field of housing and urban economics. Subjects include the demand for housing, the theory of housing, location, property values and race, slums and urban renewal, public policy, and the urban environment. © 1969, 450 pages, hardbound \$6.50

Also of Interest

The Age of the Economist

By Daniel R. Fusfeld, The University of Michigan

This widely acclaimed exposition of the major trends in economic thought presents the ideas of economists in relation to the political and philosophical ideas and circumstances of their ages. Fusfeld's application of the "sociology of knowledge" approach emphasizes the significance of economics for public policy issues, ideological disputes, and social philosophies. "Indeed, one of Fusfeld's singular merits as a writer is his capacity to sum up in one paragraph—sometimes in one inspired sentence—the seminal idea of a minor figure."
—*Saturday Review* © 1966, 157 pages, hardbound \$4.75, softbound \$3.10



Scott, Foresman College Division

Glenview, Ill. Palo Alto, Ca. Tucker, Ga. Oakland, N.J. Dallas, Tx.

Forthcoming in 1970 from Macmillan

ECONOMIC DYNAMICS, Third Edition

By William J. Baumol, Princeton University

In this third edition of his well-known book, Professor Baumol discusses the newer theories of economic growth.

1970, 472 pages, \$10.95

MONETARY AND FISCAL POLICY IN A GROWING ECONOMY

By Duncan K. Foley, Massachusetts Institute of Technology

Macmillan Series in Economics

This book synthesizes recent advances in monetary theory and relates to the neo-classical model of economic growth.

1970, approx. 288 pages, prob. \$9.95

MATHEMATICAL THEORIES OF ECONOMIC GROWTH

By Edwin Burmeister, University of Pennsylvania, and
Rodney Dobell, University of Toronto, with a Foreword by
Robert M. Solow, Massachusetts Institute of Technology

Macmillan Series in Economics

Mathematical Theories of Economic Growth is the first book to treat this topic at an advanced level of economic and mathematical sophistication.

1970, approx. 448 pages, prob. \$12.95

STATIC DEMAND THEORY

By Donald W. Katzner, Wharton School of Finance and Commerce,
University of Pennsylvania

Macmillan Series in Economics

Using modern mathematical analysis, Professor Katzner collects, for the first time in a systematic and comprehensive manner, the results of static demand theory.

1970, approx. 256 pages, prob. \$9.95

Write to the Faculty Service Desk for examination copies.

THE MACMILLAN COMPANY

866 Third Avenue

New York, New York 10022

In Canada, write to Collier-Macmillan Canada, Ltd.,
1125B Leslie Street, Don Mills, Ontario

Praeger College Department

SEEDS OF CHANGE

The Green Revolution and Development in the 1970's

LESTER R. BROWN. Foreword by **EUGENE R. BLACK.** The first full account of an astonishing agricultural breakthrough—the development of new strains of “miracle” rice and wheat which have brought a victory over hunger and malnutrition within man's grasp. “Everyone concerned with development should read Lester Brown's admirable study . . . which gives the best account available of the full scale of the Green Revolution and of its impact on every aspect of development strategy.”—**BARBARA WARD.**

224 pp. \$6.95 Cloth, \$2.50 Paper

21 POPULAR ECONOMIC FALLACIES

EZRA J. MISHAN. “The perfect book for anyone wishing to start the study of economics. . . . First, it deals with twenty-one problems that really matter, ranging from rent control to economic growth. Second, economic analysis is introduced and used at the same time. Third, it contains neither arid technicalities nor boring description. Fourth, it is well written. Fifth, and most important, it makes the reader want to argue. Pedagogically, therefore, it is ideal.”—*The Economist.*

April. 224 pp. \$7.95 Cloth

TECHNOLOGY AND GROWTH The Price We Pay

EZRA J. MISHAN. Professor Mishan, author of the widely acclaimed *The Costs of Economic Growth*, now focuses on our current obsession with economic growth and points to inept social policies and to the high costs of unchecked technological development. Mishan shows that, although consumer gadgetry is increasing, the choice of environment—a vital factor for human happiness and welfare—is rapidly vanishing.

April. 197 pp. \$7.95 Cloth

PARTNERS IN DEVELOPMENT

Report of the Commission on International Development

LESTER B. PEARSON. “I am convinced that it will become one of the most important documents of the twentieth century.”

—**PRIME MINISTER HAROLD WILSON.** “Begins with the rationale for aid and a review of the recent history of international cooperation for development . . . and concludes with recommendations for trade policy, private foreign investment, population control, and an international framework for development. . . . Thorough and sophisticated.”—*Library Journal.*

352 pp. \$7.00 Cloth, \$2.95 Paper

REGIONAL ECONOMICS

Location Theory, Urban Structure, Regional Change

HARRY W. RICHARDSON. “Rigorous, well-organized, tightly written and shows a good knowledge of the literature. . . . Nothing on the market can compete with it.”—**PROFESSOR JOHN NIEDERCORN.** Richardson's text is the first work to survey theoretical developments in this new and topical branch of economics in a single comprehensive volume.

464 pp. \$10.00 Cloth, \$4.95 Paper

REVENUE-SHARING Crutch or Catalyst for State and Local Governments?

CONGRESSMAN HENRY S. REUSS. Foreword by **HUBERT H. HUMPHREY.** Congressman Reuss of Wisconsin takes a hard, informative look at the desperate lack of funds that beset government at the city and state levels and proposes a dramatic solution: the sharing of federal revenues with states and localities to meet their enormous obligations. Reuss also examines President Nixon's recent revenue-sharing proposal. 184 pp. \$6.50 Cloth

PRAEGER PUBLISHERS

111 Fourth Ave. • New York 10003

Please mention *THE AMERICAN ECONOMIC REVIEW* When Writing to Advertisers



Money and the Economy **SECOND EDITION**

JOHN J. KLEIN, Georgia State University

Revised, expanded, and brought up to date for this Second Edition, *Money and the Economy* offers a basic but rigorous exposition of banking, theoretical analysis, and economic policy. Taking as his central theme the importance of money in the economy, Professor Klein treats clearly and concisely the four main areas of money and banking: the private financial institutions that determine the extent of our monetary resources; the influence of government on our monetary resources; the development of monetary theory; and alternative economic policies. Special features of this new edition include an expanded discussion of banking operations, an expanded treatment of international economics, and full coverage of such important new developments as the theory of the money supply and the utilization of Regulation Q as an instrument of economic policy.

488 pages. \$9.95 Just published

The Neurotic Trillionaire **A SURVEY OF MR. NIXON'S AMERICA**

NORMAN MACRAE, The Economist

First published in the May 10, 1969, issue of *The Economist*, this work is an urbane, cogent analysis of the development of economic policies during the Kennedy, Johnson, and Nixon administrations. Mr. Macrae provides a clear overview of the triumphs and frustrations of the designers of our economic policy; discusses the problems in political economics currently facing the federal, state, and local governments; analyzes the strengths and weaknesses of the Keynesian theory in action; and evaluates the current role of macroeconomic policy in the economy. In a brief Preface, [written in November 1969], the author brings his survey up to date and offers some new insights into possible developments during the 1970's.

Paperbound. 128 pages, \$1.95 (probable) Publication: April 1970

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Macroeconomic Analysis **SECOND EDITION**

EDWARD SHAPIRO, University of Toledo

Intended as a basic textbook for use on the intermediate level, *Macroeconomic Analysis*, Second Edition, continues to offer a thorough, balanced coverage of the essentials of macroeconomic theory. It progresses smoothly and gradually in complexity of subject matter, developing a rigorous analysis through a carefully organized and readable presentation. The Second Edition includes a number of new features that bring the book up to date and make it more accessible to intermediate students—among them, an increased emphasis on policy and growth, including treatment of neo-classical growth theory, Friedman's monetary policy, and the full-employment budget surplus; a completely new chapter on the business cycle theory; a full discussion of wage and price guidelines and their history; and a new analysis of the permanent income hypothesis. In addition, the material on national income accounting, while retaining the same broad coverage as in the earlier edition, has been reduced in length by one-third.

591 pages. \$10.50 Just published

To accompany the textbook

MACROECONOMIC ANALYSIS **A Student Workbook**

BRUCE GENSEMER, Kenyon College

EDWARD SHAPIRO, University of Toledo

Paperbound. 240 pages, \$3.50 (probable) Publication: May 1970

MACROECONOMIC ANALYSIS **Selected Readings**

Edited by EDWARD SHAPIRO, University of Toledo

Paperbound. 496 pages, \$4.95 (probable) Publication: March 1970



HARCOURT, BRACE & WORLD, INC.

New York / Chicago / San Francisco / Atlanta

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

The Poor Ye Need Not Have:

Lessons from the War on Poverty

by Robert Levine

"This is a good, solid book, as I would expect from Bob Levine. It is informed with far-sighted perspective, solid reconstruction of debates, and a thoroughly sensible review of OEO to date." — Adam Yarmolinsky, Harvard University \$7.95

The Economics of Insurgency in the Mekong Delta of Vietnam

by Robert L. Sansom

This is the most complete work to date dealing with the economy of the Mekong Delta. The results of the study have already produced striking changes in U.S. economic policy in South Vietnam, particularly in the essential areas of economic assistance and land reform. \$12.50

Trade, Tariffs, and Growth

by Jagdish Bhagwati

Professor Bhagwati has brought together his most important theoretical writings on international economics. A major contribution to the pure theory of international trade — his *Economic Journal* survey of the subject — is reprinted and provided with an addendum which brings it up to date. In addition there are papers on propositions relating to gains from trade and papers on tariffs, quotas, and subsidies that cover both "positive" and welfare aspects of trade theory. Reprinted here are his well-known papers on immiserizing growth, including a recent generalization, and on the theory of optimal policy intervention under domestic distortions. Four important essays on growth and development round out the volume. \$12.50

The International Corporation:

A Symposium

edited by Charles P. Kindleberger

This volume brings together papers by a distinguished group of scholars on various aspects of the international corporation: theoretical, financial, technological, legal, and political. The book also includes studies of three world industries — petroleum, automobile, and banking — and of three world regions where the international corporation has considerable predominance: Australia, Japan, and Latin America. These wide-ranging forays are brought to conclusion by a summary article on the future of the multinational enterprise, which treats the theory of direct investment and attitudes of home and host countries. \$15.00

Stochastic Models of Buying Behavior

by William F. Massy, David B. Montgomery, and Donald G. Morrison

This is the first book on stochastic models of consumer buying behavior, and indeed one of a relative few books dealing with original contribution to research for marketing management (as opposed to textbooks) published to date. The purpose of the book is to present an integrated treatment of the authors' research on the title subject. Including the 15% or so of material discussing the work of others in the field, and one chapter on statistical methodology, the book covers almost all aspects of the field of stochastic process models of buyer behavior as it is known today. \$17.50

Technical Resource Management:

Quantitative Methods

by Marvin J. Cetron, Raymond Isenson, Jacob N. Johnson, Ambrose B. Nutt, Howard A. Wells

Traditional research management methods, the authors contend, are too unsystematic to continue to control the funding and allocation of men, money, and materials in research. Moreover, they can now be supplanted: over the past decade a wide range of planning tools have been developed and tested; the authors believe they work; and this book gives an overview of the most significant resource allocation techniques now being used in government and industry. \$10.00

All prices tentative

The MIT Press

Cambridge, Massachusetts 02142

FROM
CONSTRAINTS
TO
OPTIONS

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Important Titles in Economics from Macmillan

COMPUTER PROBLEM KIT FOR ECONOMICS

By George I. Treyz, University of Massachusetts

Macmillan Series in Economics

Used effectively with beginning students, this kit shows how to solve almost any macro-economic model by adding or replacing cards in the prepunched program.

1969, \$2.95

Instructor's Manual *gratis*

ECONOMETRIC GAMING:

A Kit for Computer Analysis of Macroeconomic Models

By L. R. Klein and M. K. Evans, both, University of Pennsylvania, and M. J. Hartley, Duke University

Macmillan Series in Economics

This realistic programmed model of the United States economy allows students and professionals to study techniques of forecasting and determine the effects of alternative economic policies and exogenous forces.

1969, \$4.95

THE INTERNATIONAL ECONOMY

Fourth Edition

By P. T. Ellsworth, Visiting Professor, University of California, Riverside, in collaboration with J. Clark Leith, University of Western Ontario

Now available in a fourth edition, this long-popular text offers a balanced treatment of institutions and analysis, with increased emphasis on modern economic theory.

1969, 594 pages, \$9.95

MONEY, WEALTH AND ECONOMIC THEORY

By Boris P. Pesek and Thomas T. Saving

"This is an excellent work showing that a theory of money and prices may be built on a general theory of value. The authors summarize and incorporate into the analysis much of the recent literature on monetary theory and 'wealth effects.'"—*Choice*

1967, 448 pages, \$10.25

MONEY AND BANKING:

Analysis and Policy, Second Edition

By Charles R. Whittlesey, Emeritus, Arthur M. Freedman, and Edward E. Herman, all, University of Pennsylvania

The second edition of this widely respected text offers expanded coverage of monetary analysis and policy.

1968, 617 pages, \$8.95

Sent on approval only

THE MACMILLAN COMPANY, 866 Third Avenue, New York, N.Y. 10022

In Canada, write to Collier Macmillan Canada, Ltd.,
1125B Leslie Street, Don Mills, Ontario

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

coming this **spring . . .**

INTRODUCTION TO SCIENTIFIC COMPUTING

GABRIEL A. PALL, I.B.M. May 1970, 650 pp., illus., \$12.50 (tent.)

AGRICULTURE, ECONOMICS, AND GROWTH, 2nd Edition

MILTON M. SNODGRASS, California State Polytechnic College; and
LUTHER T. WALLACE, University of California, Berkeley April
1970, 500 pp., illus., \$8.50 (tent.)

HOW available . . .

CASES IN COMMERCIAL BANKING

EDWARD W. REED, National Bank of Oregon and DON L. WOODLAND, Louisiana State University 250 pp., paper, \$3.95 (tent.)

THE ORGANIZATION MAKERS

ORVIS COLLINS, State University of New York at Buffalo; and DAVID G. MOORE, New York State School of Industrial and Labor Relations (Cornell University) 250 pp., \$5.95 (tent.)

SALES STRATEGY: Cases and Readings

Edited by ROBERT F. GWINNER and EDWARD M. SMITH, both of the University of Alabama. 558 pp., illus., paper, \$5.95

DIMENSIONS OF COMMUNICATION

Edited by LEE RICHARDSON, The Louisiana State University. 488 pp., illus., paper, \$5.95

GOVERNMENTAL PROMOTION AND REGULATION OF BUSINESS

ELMER E. SMEAD, Dartmouth College. 582 pp., \$8.95

READINGS IN FINANCIAL MARKETS AND INSTITUTIONS

Edited by J. VAN FENSTERMAKER, University of Mississippi. 477 pp., illus., paper, \$5.95

CAPITAL ALLOCATION THEORY: The Study of Investment Decisions

GERALD A. FLEISCHER, University of Southern California. 291 pp., illus., \$6.95

PRODUCTS AND MARKETS

WILLIAM H. REYNOLDS, Wayne State University. 244 pp., illus., paper, \$3.50

READINGS IN GLOBAL MARKETING MANAGEMENT

Edited by ROBERT PATTY and HARVEY L. VREDENBURG, both of Colorado State University. 504 pp., illus., paper, \$5.25

PUBLIC RELATIONS SURVEY RESEARCH

EDWARD ROBINSON, Boston University. 282 pp., illus., \$6.50

THE SOVIET ECONOMY, 2nd Edition

Edited by HARRY G. SHAFFER, The University of Kansas. 628 pp., illus., \$5.95

COMPARATIVE ECONOMIC SYSTEMS

JAN S. PRYBYLA, The Pennsylvania State University. 551 pp., illus., \$8.50

MACROECONOMIC THEORY: Selected Readings

Edited by HAROLD R. WILLIAMS and JOHN D. HUFFNAGLE, both of Kent State University. 527 pp., illus., paper, \$5.95

UNDERSTANDING QUANTITATIVE ANALYSIS

CHESTER R. WASSON, Northern Illinois University. 264 pp., paper, \$3.50



APPLETON-CENTURY-CROFTS

Educational Division

MEREDITH CORPORATION

440 Park Avenue South, New York 10016

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Announcing the first volumes of NEW DIRECTIONS IN MANAGEMENT AND ECONOMICS

Reflecting the latest trends in management and economic research, this new series of Praeger books offers in-depth analyses of key problems, geared for more effective, more sophisticated decision-making. The books selected for New Directions in Management and Economics are designed to speed the implementation of new ideas and the use of the most advanced techniques.

First volumes in the series:

PROFESSIONAL DECISION-MAKING
A Procedure for Evaluating Complex Alternatives. James R. Miller III. A totally new and tested procedure for minimizing risk and avoiding costly errors, based on the author's intensive research. He shows how alternatives that include "hidden" or subjective values can be accurately defined, measured, and systematically taken into account in the decision process. Includes three major case studies. References. Appendixes. Index. April, 1970. \$12.50

ECONOMIC FORECASTING. Herman O. Stekler. A valuable book for forecasters whose work depends on a sophisticated evaluation of economic trends. The author analyzes the procedures that will provide the best estimates of key economic variables for determining economic trends. He discusses major new forecasting techniques—their strengths and weaknesses—and clearly points out the pitfalls in forecasting—the errors that can arise from inaccurate data selection and use. Appendixes. Index. April, 1970. \$11.50

THE IMPACT OF COMPUTERS ON ORGANIZATIONS. Thomas L. Whisler. A carefully researched study that analyzes the actual effects of introducing computers in organizations and

shows how computers can be used more effectively to perform organizational tasks. The author's findings are based on a long-range study in which the effects of the computers on several companies were examined by a research team actually working in these companies. Appendixes. Index. \$11.00

Forthcoming volumes:

LONG-RANGE SALES FORECASTING

By J. Scott Armstrong

PORTFOLIO EVALUATION

By Michael C. Jensen

MODERN CAPITAL THEORY

Edited by Michael C. Jensen

CONGLOMERATE MERGERS

Edited by Paul W. MacAvoy

THE COMPUTER INDUSTRY

By Ralph E. Miller

CASH MANAGEMENT AND THE DEMAND FOR MONEY

By Daniel Orr

DEVELOPING EFFECTIVE MANAGERS

A New Approach to
Management Education

By Reginald W. Revans

INVESTMENT ANALYSIS

New Analytic Techniques

By J. Peter Williamson

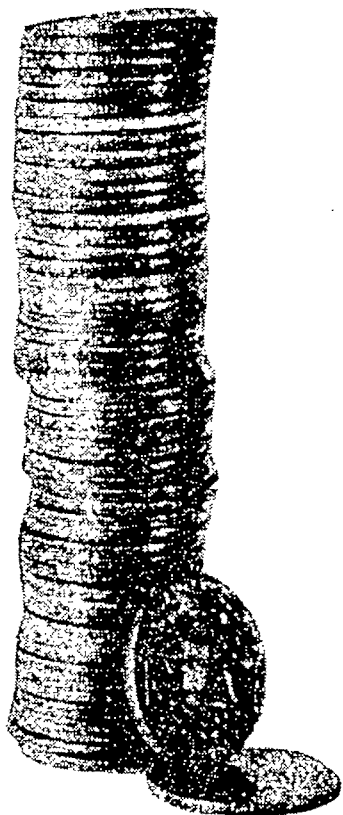
Write for descriptive brochure to:

PRAEGER PUBLISHERS

111 Fourth Avenue • New York, N.Y. 10003

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Goodyear texts



ECONOMICS AND THE PRIVATE INTEREST

Richard T. Gill, Master of Leverett House and Lecturer on Economics, Harvard University

...provides an introduction to microeconomics built around the fundamental question of the relation of private interest to social and economic welfare.

Focusing on the basic problems of interdependence and efficiency, Dr. Gill presents a detailed discussion of the workings of a competitive price system. With this theoretical background established, the text then covers areas of modern economic life where private interest and social welfare may fail to coincide. March, 1970 288 pp.

Also available: **STUDENT GUIDE TO ECONOMICS AND THE PRIVATE INTEREST: MICROECONOMIC PROBLEMS**, Robert P. Inman and Samuel A. Rea, Jr., both at Harvard University

ECONOMICS AND THE PUBLIC INTEREST

Richard T. Gill, Harvard University

...comprehensive coverage of the great economic issues of our day

Economic tools are applied to the problems and issues of immediate interest: inflation, unemployment, balance of payment problems, foreign aid, public vs. private decision-making with respect to allocation of resources, poverty, and the problems of affluence.

These present-day economic issues are also related to the history of economic thought through the examination of the writings of Smith, Ricardo, Malthus, Say, Marx, Keynes, and Galbraith.

1968 306 pp. paper \$4.95 (22729-8) / cloth \$5.95 (22730-6)

Also available: **STUDENT GUIDE TO ECONOMICS AND THE PUBLIC INTEREST**, by Kenneth M. Deitch, Harvard University
116 pp. paper \$2.95 (22731-4)

INFORMATION PROCESSING

Frank J. Clark, Genesee Community College

...ideal for preparing students for the practical application of data processing

This text gives the student a broad foundation that will enable him to learn whatever new systems, concepts, or new generation of hardware he may encounter. In addition, the book provides the student with problem-solving skills so he can be immediately productive on his first job.

An Instructor's Manual is available on adoption.

January, 1970 320 pp. \$8.95

Also available: **WORKBOOK TO INFORMATION PROCESSING**, by Frank J. Clark. The workbook contains examples of business EDP problems and provides daily assignments for the student.
January, 1970 80 pp. \$4.50 (48403-0)

PRICE THEORY AND APPLICATIONS IN BUSINESS ADMINISTRATION

W. David Maxwell, Indiana University

...a text for the first course in microeconomics required of candidates for the master's degree in business administration

Professor Maxwell has tailored his new text to the needs of the advanced undergraduate or beginning graduate student in business administration. Nonetheless, the book treats price theory as a system prior to application so the student gains a solid foundation in the subject. Both specific applications and theoretical extensions are made in several areas of business administration.

January, 1970 270 pp. \$10.00 (69958-7)

ECONOMIC PROBLEMS AND POLICIES

John M. Kuhlman, University of Missouri

... explains the elementary concepts in economics and presents a variety of current economic problems

Designed for a one-semester course, this basic textbook shows how economics can be used to solve current economic problems. The first section presents the basic tools of economic analysis; the second is a collection of readings that clearly illustrates the ways economists use these basic analytical tools. The readings, which come primarily from Congressional documents, deal with problems of current concern.

1969 368 pp. paper \$5.95 (22701-7)

PUBLIC POLICY TOWARD MERGERS

J. Fred Weston and Sam Peltzman, both of UCLA

... provides a valuable compendium of essays on public policy toward mergers by outstanding authorities in the field

The editors summarize important and related empirical research including leading court decisions on anti-trust. In addition, they raise many issues on unresolved policy questions, making the book a useful guideline for future research. It can be used in courses on economic principles, industrial economics, business and society, and antitrust economics.

1969 256 pp. paper \$4.95 (73826-0)

READINGS IN INTERNATIONAL FINANCIAL MANAGEMENT

Ed. by Arthur T. Stonehill, Oregon State University

... makes available in one volume many of the recently published articles presently being used in courses of international financial management

Dr. Stonehill has carefully chosen readings that challenge and extend selected topics in corporate finance. Readings are included from BUSINESS INTERNATIONAL and the National Industrial Conference Board. An ideal supplementary text, this book can be easily integrated into courses in corporate finance, international business finance, and international finance.

January, 1970 263 pp. \$4.95 (75754-2)

PROGRAM BUDGETING AND BENEFIT COST ANALYSIS: CASES, TEXT AND READINGS

Harley H. Hinrichs, U.S. Naval Academy and University of Maryland; Graeme M. Taylor, University of Maryland

... presents the principles of government decision-making under the planning-programming-budgeting system

The text accomplishes this through the most current readings in the field and case studies. The cases represent the first published collection of teaching cases on public expenditure analysis and management. They have all been class-tested in training programs for government officials at all levels and in graduate seminars in university courses.

1969 464 pp. \$11.95 (72981-4)

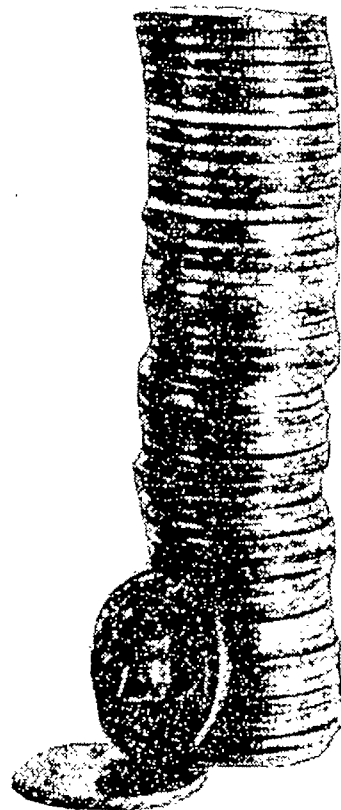
For more information, write:

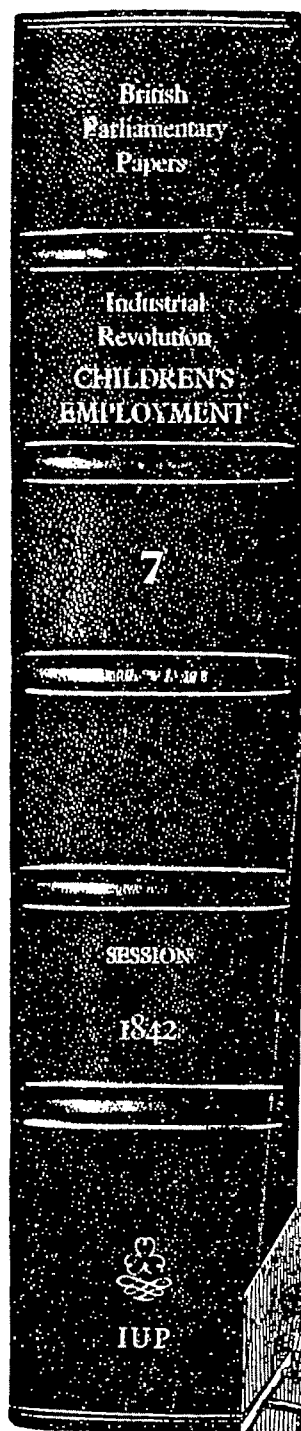
GOODYEAR PUBLISHING COMPANY, INC.

15115 Sunset Blvd.

Pacific Palisades, California 90272

cover important aspects of Economics





Children's Employment

15 Volumes. Industrial Revolution Subject Set of
The Irish University Press Series of The British Parliamentary Papers

Perhaps the most sordid chapter in the history of labor concerns the abuse inflicted on the very young during the industrial revolution. The cost of industrial progress was reckoned in terms of the most appalling human misery.

Irish University Press has drawn the original, rich, source-material—extensive records of government investigations and studies, the evidence of witnesses, including the witness of the children themselves—into a 15-volume subject set of the British Parliamentary Papers, unquestionably the most valuable and important research tool for students and teachers of the social sciences.

The 15 volumes of documentation cover over one-half century of investigations from the 1816 report of a select parliamentary committee established to investigate the four "radical" proposals of the great social reformer, Robert Owen, to the 1867 reports on the safety provisions, accommodations and sanitary regulations of those trades which still were not regulated by some law.

The history of child labor is documented in over 10,000 pages of text, illustrations, charts, diagrams and tables. The volumes offer a complete analysis of the ages and sex of employed children, hours spent working each day, education, factory inspections, occupational diseases, physical mistreatment, health conditions and hazards, holidays, mortality rates, housing, wages, accidents, food, and sanitary conditions. The industries investigated included mining, textile manufacturing, paper mills, tobacco, iron and metal goods, glass and potteries manufacturers, printing plants, bleaching and dyeing, and the clothing trades.

Children's Employment is one of the 75 subject sets in the Irish University Press 1000-volume Series of the British Parliamentary Papers, making available to student and scholar researchers an almost inexhaustible font of contemporary knowledge of an era when humanitarianism was beginning to mold legislation. *Children's Employment* is available as a subject set or as part of the entire British Parliamentary Papers Series.

A free descriptive 16-page catalog detailing each of the 15 volumes in the *Children's Employment* subject set will be sent to you. Return the coupon or write today to receive your copy.

IRISH UNIVERSITY PRESS, Inc. 2 Holland Ave White Plains, N.Y. 10603

- ☐ Send, free, descriptive 16-page catalog on the *Children's Employment* subject set.
- ☐ Send, free, additional information on the entire Irish University Press series of the British Parliamentary Papers.
- ☐ I want to discuss this project with your sales representative.

Name _____

Institution _____

Address _____

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

WORKERS AND WAGES IN AN URBAN LABOR MARKET

Albert Rees and George P. Shultz

In this book two of the most distinguished labor economists in the country analyze the factors that determine wage differentials within occupations. Their findings are based on a five-year study of the information obtained primarily from personnel records of five thousand workers in seventy-five establishments. This is one of the first labor market studies to cover a major metropolitan area and to deal with professional workers, clerical workers, and blue-collar workers using the same methodology. The authors demonstrate that wage differentials are related to a number of personal characteristics including seniority, schooling, previous experience, sex, and race. George P. Shultz is currently Secretary of Labor. *Studies in Business and Society* series.

1970 LC:75-110114 244 pages \$10.00

PIONEERS FOR PROFIT

Foreign Entrepreneurship and Russian Industrialization, 1885-1913

John P. McKay

Foreign investment increased from 17 percent of the capital of industrial corporations in Imperial Russia in 1880 to 47 percent in 1914, coinciding with the rapid development of Russian industrialization before World War I. John McKay's study is the first detailed analysis of the impact of foreign enterprise on Russian industry during this period. Foreign investors, he notes, saw their more advanced industrial techniques as the key to large profits and built plants which far surpassed usual Russian practice. While only a few of these pioneers actually made their large profits, they made a positive contribution to Russia's industrial development. McKay's conclusions will be significant for historians, economists, and those interested in the development of modern industrial society.

1970 LC:79-103932 428 pages \$11.50

CHICAGO

THE GATT

Law and International Economic Organization

Kenneth W. Dam

A case study in international organization, this book offers a legal analysis with economic and political implications. The subject of the study is the General Agreement on Tariffs and Trade, or the "GATT." Dealing day by day with the problems of international trade in their many dimensions, the GATT has become the principal international trade institution of the postwar world. The author is interested in the role of rules in international life—particularly procedural rules that can be used to influence sovereign nations. He argues that only when law is understood not as a set of prohibitions but as a set of procedures to be used in reaching agreements and accommodations can it become an effective force in world affairs.

1970 LC:75-93088 544 pages \$15.00

The University of Chicago Press, Chicago 60637

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

*Since publication
in February, 1969,
more than 140 colleges
and universities have adopted
the SECOND EDITION of*

ECONOMICS

by Richard G. Lipsey, UNIVERSITY OF ESSEX
and Peter O. Steiner, THE UNIVERSITY OF MICHIGAN

Lipsey and Steiner's **ECONOMICS** has, in its two editions, reflected the profound change in method and point of view that the discipline has undergone during the past two decades. The approach, the method, and the structure of this book have been planned to teach economics as it is practiced—as the genuine social science it has become. The format of the **SECOND EDITION** is entirely new. It includes expanded treatment of *Macroeconomics* (rewritten, reorganized, and more than doubled in size, with the number and variety of examples substantially increased); *Policy* (major policy discussions and materials expanded and better integrated); *Demand Theory* (now includes marginal utility theory, indifference theory, and revealed preference theory); *Public Policy Toward Monopoly* (with a discussion of public utility regulation); *Fiscal Policy*; and *Growth and Underdeveloped Economies* (rewritten to incorporate much up-to-date comparative data). There are many new chapters, including *Price Theory in Action*, *Monetary Policy*, and *Comparative Economic Systems*. 845 pp.; \$9.95

To accompany the text: Study Guide by Dorothy Forbush and Dascomb R. Forbush (341 pp.; \$3.50) and Instructor's Manual by Lipsey and Steiner.

HARPER & ROW, Publishers,
49 East 33d Street, New York 10016



Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

*New and notable
Harper & Row
texts of interest*

THE ECONOMICS OF MONEY AND BANKING, Fifth Edition

LESTER V. CHANDLER, Princeton University

1969. 576 pp.; \$9.95 / Student Guide—Money, Banking, and Monetary Theory: Problems and Concepts (by Harold R. Williams and Henry W. Woudenberg): 224 pp.; \$8.95

MONEY, BANKING, AND MONETARY POLICY

Readings in Domestic and International Policy

HAROLD R. WILLIAMS and HENRY W. WOUDENBERG, Kent State University

June, 1970. Tentative: \$5.95 paper

ECONOMETRICS

Statistical Foundations and Applications

PHOEBUS J. DHRYMES, University of Pennsylvania

January, 1970. 584 pp. + index; \$14.95

MATHEMATICAL ANALYSIS

Business and Economic Applications

JEAN E. DRAPER, University of Arizona, and JANE S. KLINGMAN, Xavier University

1967. 568 pp.; \$13.50

MACROECONOMIC ACTIVITY

Theory, Forecasting, and Control—An Econometric Approach

MICHAEL K. EVANS, University of Pennsylvania

1969. 627 pp.; \$15.95

ECONOMIC STATISTICS AND ECONOMETRICS

An Introduction to Quantitative Economics

EDWARD J. KANE, Boston College

1968. 437 pp.; \$11.95

COMPARATIVE ECONOMIC SYSTEMS, Eighth Edition

WILLIAM N. LOUCKS and WILLIAM G. WHITNEY, University of Pennsylvania

1969. 582 pp.; \$9.95

INTERNATIONAL ECONOMICS

BO SÖDERSTEN, University of Lund, Sweden

March, 1970. 556 pp.; \$13.95

STATISTICS, AN INTRODUCTORY ANALYSIS, Second Edition

TARO YAMANE, New York University

1967. 919 pp.; \$11.95 / Statistical Tables: \$37.50 per pkg. of 50 / Problems Book: \$1.95

Send for our 1970 economics catalogue / #36-00475

HARPER & ROW, Publishers,

49 East 33d Street, New York 10016



1817

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

from an interview with the author

DANIEL B. SUITS

HOW DID YOUR INTRODUCTORY TEXTBOOK COME TO BE WRITTEN?

In recent years, conversations with students in my principles course led me to conclude that traditional presentation of introductory economics is inadequate for today's students. The difficulty, I feel, is that there is too little direct contact between the material in the course and what is going on in the student's world. I began to write short chapters that I thought would fill gaps, but I soon found that what is really needed is an entirely different approach—one that begins with the world as it can be observed, that allows economic principles to emerge from our environment, and that permits the student in reading his newspaper to apply what he has learned. To do this requires not only a different kind of textbook, but also a set of practical problem materials for the student to work out himself and a way of bringing the daily newspaper into class discussion.

BECAUSE OF YOUR REPUTATION AS AN ECONOMETRICIAN, WON'T INSTRUCTORS EXPECT YOUR BOOK TO BE HEAVILY QUANTITATIVE? IS THE BOOK HIGHLY MATHEMATICAL? No. On the contrary, if anything, it is probably less mathematical than most principles texts. Once the student's sights are directed toward what's going on in the world, he can easily follow the economic discussion because he knows what the material is directed toward. There are inevitably a few formulas, but I believe that there are far fewer in this book than there are in most books.

Principles of Economics

Daniel B. Suits, The University of Michigan. A new approach to introductory economics, this book has been six years in development and thoroughly class-tested. The determinants of productive capacity are examined via the live setting of world poverty, and demographic and economic factors are integrated into a unified dynamic analysis. Special chapters deal with poverty in the United States and with the economics of environmental pollution. By studying the actual behavior of the economy, students are led to do their own forecasting. Differences in pricing behavior are examined both theoretically and empirically. Monetary and fiscal policy are integrated into a unified theory of stabilization policy that shows the two as complements rather than rivals. One hundred years of U.S. economic history are surveyed as a case study of economic development. An Instructor's Manual and a Student Guide (tentative: \$3.95) are available. MARCH, 1970. 543 PAGES; \$9.95

HARPER & ROW, Publishers

49 e. 33d street, new york 10016



1817

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

The American Economic Review

PAPERS AND PROCEEDINGS

OF THE

Eighty-second Annual Meeting

OF THE

AMERICAN ECONOMIC ASSOCIATION

New York, New York, December 28-30, 1969

Edited by the Secretary of the Association

MAY 1970

AMERICAN ECONOMIC ASSOCIATION
Organized at Saratoga, New York, September 9, 1885

PAST OFFICERS

Presidents

1. FRANCIS A. WALKER,* M.I.T., 1886-92
2. CHARLES F. DUNBAR,* Harvard, 1893
3. JOHN B. CLARK,* Columbia, 1894-95
4. HENRY C. ADAMS,* Michigan, 1896-97
5. ARTHUR T. HADLEY,* Yale, 1898-99
6. RICHARD T. ELY,* Wisconsin, 1900-01
7. EDWIN R. A. SELIGMAN,* Columbia, 1902-03
8. FRANK W. TAUSSIG,* Harvard, 1904-05
9. JEREMIAH W. JENKS,* Cornell, 1906-07
10. SIMON N. PATTEN,* Pennsylvania, 1908
11. DAVIS R. DEWEY,* M.I.T., 1909
12. EDMUND J. JAMES,* Illinois, 1910
13. HENRY W. FARNAM,* Yale, 1911
14. FRANK A. FETTER,* Princeton, 1912
15. DAVID KINLEY,* Illinois, 1913
16. JOHN H. GRAY,* Minnesota, 1914
17. WALTER F. WILLCOX,* Cornell, 1915
18. THOMAS N. CARVER,* Harvard, 1916
19. JOHN R. COMMONS,* Wisconsin, 1917
20. IRVING FISHER,* Yale, 1918
21. HENRY B. GARDNER,* Brown, 1919
22. HERBERT J. DAVENPORT,* Cornell, 1920
23. JACOB H. HOLLANDER,* Johns Hopkins, 1921
24. HENRY R. SEAGER,* Columbia, 1922
25. CARL C. PLEHN,* California, 1923
26. WESLEY C. MITCHELL,* Columbia, 1924
27. ALLYN A. YOUNG,* Harvard, 1925
28. EDWIN W. KEMMERER,* Princeton, 1926
29. THOMAS S. ADAMS,* Yale, 1927
30. FRED M. TAYLOR,* Michigan, 1928
31. EDWIN F. GAY,* Harvard, 1929
32. MATTHEW B. HAMMOND,* Ohio State, 1930
33. ERNEST L. BOGART,* Illinois, 1931
34. GEORGE E. BARNETT,* Johns Hopkins, 1932
35. WILLIAM Z. RIPLEY,* Harvard, 1933
36. HARRY A. MILLIS,* Chicago, 1934
37. JOHN M. CLARK,* Columbia, 1935
38. ALVIN S. JOHNSON, New School, 1936
39. OLIVER M. W. SPRAGUE,* Harvard, 1937
40. ALVIN H. HANSEN, Harvard, 1938
41. JACOB VINER, Chicago, 1939
42. FREDERICK C. MILLS,* Columbia, 1940
43. SUMNER H. SLICHTER,* Harvard, 1941
44. EDWIN G. NOURSE, Brookings, 1942
45. ALBERT B. WOLFE,* Ohio State, 1943
46. JOSEPH S. DAVIS, Stanford, 1944
47. I. L. SHARFMAN,* Michigan, 1945
48. E. A. GOLDENWEISER,* Institute for Advanced Study, 1946
49. PAUL H. DOUGLAS, Chicago, 1947

* Deceased.

OFFICERS FOR 1970

President

WASSILY LEONTIEF, Harvard University

President-elect

JAMES TOBIN, Yale University

Vice-Presidents

EVSEY D. DOMAR, Massachusetts Institute of Technology
TIBOR SCITOVSKY, Yale University

Secretary-Treasurer and Editor of Proceedings

HAROLD F. WILLIAMSON, Northwestern University

Managing Editor of the American Economic Review

GEORGE H. BORTS, Brown University

50. JOSEPH A. SCHUMPETER,* Harvard, 1948
51. HOWARD S. ELLIS, California, 1949
52. FRANK H. KNIGHT, Chicago, 1950
53. JOHN H. WILLIAMS, Harvard, 1951
54. HAROLD A. INNIS,* Toronto, 1952
55. CALVIN B. HOOVER, Duke, 1953
56. SIMON KUZNETS, Pennsylvania, 1954
57. JOHN D. BLACK,* Harvard, 1955
58. EDWIN E. WITTE,* Wisconsin, 1956
59. MORRIS A. COPELAND, Cornell, 1957
60. GEORGE W. STOCKING, Vanderbilt, 1958
61. ARTHUR F. BURNS, Columbia, 1959
62. THEODORE W. SCHULTZ, Chicago, 1960
63. PAUL A. SAMUELSON, M.I.T., 1961
64. EDWARD S. MASON, Harvard, 1962
65. GOTTFRIED HABERLER, Harvard, 1963
66. GEORGE J. STIGLER, Chicago, 1964
67. JOSEPH J. SPENGLER, Duke, 1965
68. FRITZ MACHLUP, Princeton, 1966
69. MILTON FRIEDMAN, Chicago, 1967
70. KENNETH E. BOULDING, Colorado, 1968
71. WILLIAM J. FELLNER, Yale, 1969

Secretaries

RICHARD T. ELY,* 1886-92
EDWARD A. ROSS,* 1893
JEREMIAH W. JENKS,* 1894-96
WALTER F. WILLCOX,* 1897-99

Treasurers

EDWIN R. A. SELIGMAN,* 1886-90
FREDERICK B. HAWLEY,* 1891-95
CHARLES H. HULL,* 1896-99

Secretary-Treasurers

CHARLES H. HULL,* 1900
FRANK A. FETTER,* 1901-06
WINTHROP M. DANIELS,* 1907-08
THOMAS N. CARVER,* 1909-13
ALLYN A. YOUNG,* 1914-20
RAY B. WESTERFIELD,* 1921-25
FREDERICK S. DEIBLER,* 1925-35
JAMES WASHINGTON BELL,* 1936-61

Editors

DAVIS R. DEWEY,* 1911-40
PAUL T. HOMAN, 1941-51
BERNARD F. HALEY, 1952-62
JOHN G. GURLEY, 1962-68

Editor of the Journal of Economic Literature

MARK PERLMAN, University of Pittsburgh

Elected Members of the Executive Committee

OTTO ECKSTEIN, Harvard University
ANTHONY SCOTT, University of British Columbia
MARY JEAN BOWMAN, University of Chicago
CHARLES L. SCHULTZE, Brookings Institution
ROBERT DOREMAN, Harvard University
ARNOLD C. HARBERGER, University of Chicago

Ex Officio Members

KENNETH E. BOULDING, University of Colorado
WILLIAM J. FELLNER, Yale University

Counsel

SEYMOUR J. RUBIN, Washington, D.C.

RICHARD T. ELY LECTURE

THE ECONOMICS OF PRODUCTION

By NICHOLAS GEORGESCU-ROEGEN
Vanderbilt University

For the last twenty years or so I have singled myself out among my fellow econometricians for arguing with all the means at my disposal that not every element of the economic process can be related to a number and, consequently, that this process cannot be represented in its entirety by an arithmomorphic model. At the same time, I have insisted that in our haste to mathematize economics we have often been carried away by mathematical formalism to the point of disregarding a basic requirement of science; namely, to have as clear an idea as possible about what corresponds in actuality to every piece of our symbolism. Curiously, in the home of quantity, in the natural sciences, this position does not constitute a singularity. On the contrary, essentially the same words of caution have come from many a high authority in physics—such as Max Planck or Percy William Bridgman, for example.¹ But even some engineers have raised their voices against blind symbolism. The recent remarks by a well-known British engineer are worth quoting at length:

Contrary to common belief it is sometimes easier to talk in mathematics than to talk in English; this is the reason why many scientific papers contain more mathematics than is either necessary or desirable. Contrary to common belief it is also often less precise to do so. For mathematical symbols have a tendency to conceal the physical meaning that they are intended to represent; they sometimes serve as a substitute for the arduous task of deciding what is and what is not relevant; . . . It is true that mathematics cannot lie. But it can mislead.

However, the dangers of over-indulgence in formula spinning are avoided if mathematics is treated, wherever possible, as a language into which *thoughts may only be translated after they have first been [clearly] expressed in the language of words*. The use of mathematics in this way is indeed disciplinary, helpful, and sometimes indispensable.²

The topic of this lecture—the economics of production—presents, I believe, sufficient interest by itself. But in choosing it, I have been guided also

by the fact that it may serve as a substantial illustration of the harm caused by the blind symbolism that generally characterizes a hasty mathematization.

I

What has come to be known as “the production function” is quite an old item in the economist’s analytical paraphernalia. As we may recall, it was introduced in 1894 by Wicksteed with one simple remark: “*the product being a function of the factors of production we have $P=f(a, b, c, \dots)$* .”³ This paradigm of imprecision apparently sufficed to make us accept Wicksteed’s simple symbolism as an adequate analytical representation of any production process and use it indiscriminately in every kind of situation. And as the usage of the vapid terms “input” and “output” became widespread, popular manuals came to treat the subject in an even more cavalier manner than Wicksteed’s. A typical presentation is that the production function expresses symbolically the fact that “the output of the firm depends on its inputs.”

But even consummate economists have accepted Wicksteed’s formula without any ado. They only felt that the meaning of the variables involved ought to be explained. The greater number of such authors adopt the position that the formula shows the quantities of inputs (or of factors) necessary to produce a certain quantity of output (or of product). Accordingly, all symbols in a production function,

$$(1) \quad Q = F(X, Y, Z, \dots),$$

stand for quantities.⁴ Others conceive the same function as a relation between the inputs per unit

¹ Philip H. Wicksteed, *The Co-ordination of the Laws of Distribution* (London, 1894), p. 4.

² For a small yet representative sample, see A. L. Bowley, *The Mathematical Groundwork of Economics* (Oxford, 1924), pp. 28–29; J. R. Hicks, *The Theory of Wages* (London, 1932), p. 237; E. Schneider, *Theorie der Produktion* (Vienna, 1934), p. 1; A. C. Pigou, *The Economics of Stationary States* (London, 1935), p. 142; P. A. Samuelson, *Foundations of Economic Analysis* (Cambridge, Mass., 1948), pp. 57–58; K. E. Boulding, *Economic Analysis* (3rd ed., New York, 1955), p. 585; Sune Carlson, *A Study on the Pure Theory of Production* (New York, 1956), p. 12; Ragnar Frisch, *Theory of Production* (Chicago, 1965), p. 41.

¹ Max Planck, *The New Science* (New York, 1959), p. 43, 158–59; P. W. Bridgman, *The Logic of Modern Physics* (New York, 1949), p. 50.

² Reginald O. Kapp, *Towards a Unified Cosmology* (New York, 1960), p. 111. My italics.

of time and the output per unit of time; i.e., as a relation

$$(2) \quad q = f(x, y, z, \dots),$$

in which all symbols stand for rates of flow.⁵

Curiously, no one seems to have been intrigued by the existence of these entirely distinct interpretations. Instead, many economists—including some analytical authorities—have used both definitions indifferently, sometimes even on the same page.⁶ The undeniable inference is that the economic profession considers relations (1) and (2) as two completely equivalent ways of representing any production process whatsoever. Yet behind this belief there lies an analytical imbroglia which is easily brought to light.

We need only recall that the production function is a tool associated with a static process or, to use a more explicit expression, with a steady-going process. For such a process, the following relations

$$(3) \quad Q = tq, X = tx, Y = ty, \dots$$

hold for any time interval t and for the quantities of product and of factors corresponding to that interval. With the aid of these relations and (2), relation (1) becomes

$$(4) \quad tf(x, y, z, \dots) = F(tx, ty, tz, \dots).$$

And since this relation must be true for any t , it follows, first, that f and F are one and the same function,

$$(5) \quad f(x, y, z, \dots) \equiv F(x, y, z, \dots),$$

and, second, that this function is homogeneous of the first degree. Therefore, the tacit presumption that the forms (1) and (2) are equivalent implies that the returns to scale must be constant in absolutely every production process.

Nothing, I believe, need be added to convince ourselves that this imbroglia is the direct consequence of our acceptance of Wicksteed's symbolism without first probing its validity as an analytical mirror of actuality. This conclusion raises a new and troublesome issue. Does either of the forms, (1) or (2), constitute an adequate representation of a process of production and, if so, what kind of process may be represented by it? For a start, let us try to examine it in its broad lines.

⁵ G. Stigler, *The Theory of Competitive Price* (New York, 1942), p. 109; T. C. Koopmans, "Analysis of Production as an Efficient Combination of Activities," in *Activity Analysis of Production and Allocation*, ed. T. C. Koopmans (New York, 1951), p. 35.

⁶ E.g., Frisch, *op. cit.*, p. 43.

II

Before anything else, we should note that for no other branch of economics is the concept of process as essential as for the economics of production. But, widely used though the word "process" is in sciences and philosophy, the literature seems to offer no specific definition of it. Now and then, the concept is merely associated with change. However, change is a notoriously intricate notion which has kept philosophers divided into two opposing camps: one maintaining that all is only being; the other, that all is only becoming. Obviously, science can follow neither of these teachings. Nor can it, unfortunately, embrace Hegel's dialectical synthesis that being is becoming. Analytical science must distinguish between object and event. Consequently, it must embrace the so-called "vulgar" philosophy—according to which there are both being and becoming—and cling to it to the very end. The upshot is that science must find a way to represent a process analytically.

It is obvious that, for this purpose, we must retain one point of dialectics; namely, that change and, hence, process cannot be conceived otherwise than as a relation between some entity and its complement in the absolute whole. In viewing a living tree as a process we oppose that tree to everything else—to "its other," in Hegel's terminology. Only for the absolute whole—the universe in its eternity—has change no meaning: such a totality has no complement. The notion of partial process, therefore, implies some slits cut into the absolute whole. But as a long series of thinkers, from the ancient Anaxagoras to the modern Niels Bohr, have taught us, this whole is seamless.⁷ However, in this case as in all others, analysis must proceed by some heroic simplifications and totally ignore their ultimate consequences.

The first heroic step is to divide actuality into two parts—one representing the partial process in point; the other, its environment (so to speak)—separated by a boundary consisting of an arithmomorphic void. For if the boundary would not be such a void, we would get three parts instead of two and, as is easily seen, we would be drawn into a dialectical infinite regress. So, all that exists in actuality at any moment must belong either to a process or to its environment. The basic element of the analytical picture of a process is, therefore, the boundary. No analytical boundary, no analytical process.

See Fragment 8 in J. Burnet, *Early Greek Philosophy* (4th ed., London, 1930), p. 47; Niels Bohr, *Atomic Physics and Human Knowledge* (New York, 1958), p. 10.

Now, precisely because actuality is a seamless whole we can slice it wherever we may please. And, Plato to the contrary,⁸ actuality has no joints to guide a carver. As economists we know only too well the unsettled issue of where the natural boundary of the economic process lies. Only our particular purpose in each case can guide us in drawing the boundary of a process. So, every scientist slices actuality in the way that suits best his own objective—an operation that cannot be performed without some intimate knowledge of the corresponding phenomenal domain.

An analytical boundary, as conceived here, must consist of two components. Like a frontier, one component separates the process at any time from the rest of actuality, although we must not think that this frontier is necessarily geographical in nature or rigidly determined. Witness the process of thought itself or that of an acorn growing into an oak. The second component is the duration of the process, determined by the time moments at which the process we have in mind begins and ends. Naturally, these moments must be at a finite distance; otherwise, we would not know all that has gone into the process or all that the process does. Nor must we allow them to coincide. For, to recall Whitehead's admonition,⁹ a durationless process, an event at an instant of time as a primary fact of nature, is nonsense.¹⁰ We can then choose the time scale so that the process begins at $t=0$ and ends at $t=T$, with $T>0$. T is the duration of the process, but for a production process we may prefer, instead, Marx's term: the time of production.

The next point is truly crucial: in saying that a given analytical process begins at $t=0$ and ends at $t=T$ we must take the underscored words in their strictest sense. Before $t=0$ and after $t=T$ the analytical process is out of existence. That is, in conceiving such a process we must totally abstract from it what happens outside the duration we have assigned to the process. The mental operation is clear: an analytical process must be viewed as a hyphen between one *tabula rasa* and another.

Our next problem is to arrive at an analytical description of the happening, associated with a given process. Because of the principle, "No analytical boundary, no analytical process," analysis

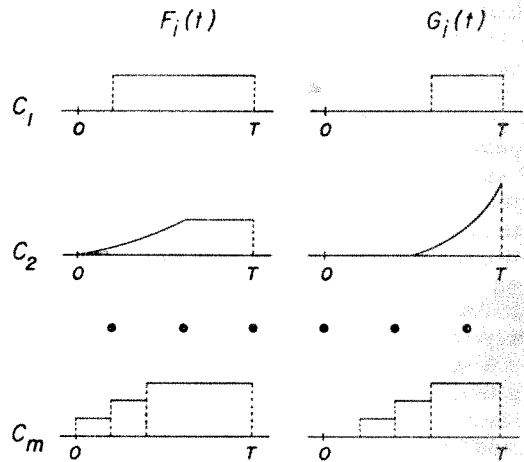


FIGURE 1

must renounce any hope of including in this description the happenings inside an analytical process. Indeed, in order to describe analytically what happens inside a process, we must divide it by a new boundary into two new processes to which the same rule will apply. And so on ad infinitum. The analytical description of a process, therefore, reduces to recording everything that crosses the boundary in either direction. In connection with this picture we can endow the terms input and output with quite precise meanings.

Analysis now needs to take another heroic step—to assume that the number of elements involved is finite and that every element is cardinally measurable (which implies that every element is a homogeneous entity). If C_1, C_2, \dots, C_m denote the distinct elements, the analytical description is complete if for every C_i we have determined two nondecreasing functions $F_i(t)$ and $G_i(t)$ over the closed interval $[0, T]$, the first function showing the cumulative input, the second the cumulative output up to time t (inclusive). Any analytical process—whether in economics or any other domain—may therefore be represented graphically by a series of curves, as in Figure 1.

In a plastic image, the coordinates of an analytical process may be likened to continuously reported data of import and export, with one important detail. Since in describing a process analytically we must begin and end with a *tabula rasa*, this hall (in which we are now gathered) must be listed both as input and as output in the process consisting of the delivery of this lecture. In the analytical approach we are not interested in how this hall came into being or in its use before or after this lecture. However, we must recognize

⁸ Plato, *Phaedrus*, p. 265.

⁹ Alfred North Whitehead, *An Enquiry Concerning the Principles of Natural Knowledge* (2nd ed., Cambridge, England, 1925), p. 2; also his *The Concept of Nature* (Cambridge, England, 1930), p. 57.

¹⁰ All this does not mean that, in the next stage of our inquiry, we cannot arrive at the excluded cases by a passage to the limit.

that, as the result of every use, the tool suffers some wear and tear, imperceptible though this may be. Similarly, in any production process the same person must be listed as a rested worker among inputs and as a tired worker among outputs. A tool, too, may go in new and come out used. But even though we recognize the rested and the tired worker as being the same man, we must treat the former as a different element from the latter. Each element of an analytical process—as we have decided—must be completely homogeneous, a condition that does not always cover sameness.

These cases, of the worker and of the tool, raise a troublesome problem for the economist. The reason is that our material of study is the commodity. We slice the economic domain into units of production and units of consumption because at the boundaries thus drawn we can catch every commodity. Drawing a boundary in a glass plant between the melted glass furnace and the rolling machines would serve none of the economist's purposes: at this time, melted glass is not a commodity.¹¹ Briefly, the economist cannot afford to abandon his commodity fetishism any more than the chemist, for example, can renounce his fetishism of the molecule.

The difficulty, which at bottom is related to qualitative change, is that even though we cannot avoid including "tired worker" and "used tool" in the list of outputs of any production process, neither category fits the usual notion of commodity. Our entire analytical edifice would collapse if we were to accept the alternative position that the aim of economic production is to produce not only the usual products but also tired workers and used tools.

A new heroic step is needed to eliminate this difficulty. It consists of the familiar, old fiction of a process in which capital is maintained constant. The fiction does raise some analytical issues, for if all tools and all workers are to be maintained at a constant level of efficiency, any production process will have to include most of the enterprises and households in the world. Factually, however, the fiction is not more, not less reasonable than that of frictionless movement in mechanics. A simple glance at the activity inside a plant or a household suffices to convince us that efforts are constantly directed not toward keeping durable goods physically self-identical (which is quite impossible), but toward maintaining them in

good working condition. And this is all that counts in production. The only factor we need neglect is the daily wear and tear of the worker. This is not too much to demand, since the worker is daily restored in the household.

The elements may now be divided into two relevant categories. In the first category we shall place those elements that appear as input and as output and are related by reason of sameness or of equality of quantity. A piece of Ricardian land, a motor, the amount of clover used as seed in growing clover seed (not clover fodder!), or a worker, illustrate this category. To elements such as these I propose to refer as funds so as to emphasize their economic invariableness. The other elements, which appear only as input or only as output, constitute the category of flows. Obviously, since the fund elements are maintained, the process may be activated again provided that the necessary inflows are still forthcoming. Labeled variously as a static or as a stationary process, or, still, as a diagram of simple reproduction, this fiction constitutes the fundamental element in the analysis of production from the classical to the hypermodern school. Reproducible, however, seems to describe the process better. The analytical picture we have thus reached is worth stressing: in a reproducible process, the fund elements are the immutable agents that transform some input flows into output flows. No picture of a process—whether static or dynamic—is complete if it does not include both categories of elements.¹² And the essential difference between these categories calls for a different representation of the fund coordinates. A flow coordinate will continue to be represented, according to the case, by the cumulative input or the cumulative output; i.e., by a quantity of some substance. Because in case of a fund the input and the output are economically the same substance, the coordinate of a fund may be represented by the difference $F_i(t) - G_i(t)$. But to maintain a convenient symmetry with the flow coordinates, we may use instead the cumulative amount of that intangible entity usually called the service of the fund.

III

In the case of a production process, the elements may be classified into some fruitful categories. The inflows that are transformed by the agents may come either from nature or from other

¹¹ That technological innovations may change this situation is evidenced by ready-mix cement and brown-and-serve rolls, for example, which only recently have become commodities.

¹² A point on which I insisted long ago: cf. my article "Aggregate Linear Production Function and Its Applications to von Neumann's Model" in *Activity Analysis of Production and Allocation*, ed. T. C. Koopmans (New York, 1951), pp. 100-01.

production processes; we shall denote them generically by (R) and (I). There also are inflows, (M), earmarked for maintenance. The output flows consist of products, (Q), and waste, (W). Finally, the funds include Ricardian land (L), capital equipment (K), and—to use Marx's very appropriate term—labor power, (H). With these notations, the analytical representation of a reproducible process is

$$(6) \{R_0^T(t), I_0^T(t), M_0^T(t), Q_0^T(t), W_0^T(t); L_0^T(t), K_0^T(t), H_0^T(t)\};$$

that is, a set of functions, which defines a point in an abstract (functional) space.

This is a far cry from the notion inherited from Wicksteed, according to which a process is represented by a point in the ordinary (Euclidean) space. The superiority of (6) over the point representation needs no elaborate argument. In (6) we have a complete set of instructions on how to set up the corresponding process. The form also reminds us continuously that a process has a duration, a time of production. Nothing is missing from it.

The difference yields an entirely new form for the production function. Since by a production function we must still understand the set of all processes that transform the same input flows into an outflow of the same product, from (6) it follows that the production function must be a relation among functions, instead of numbers. This relation, which may be written after the old pattern as

$$(7) Q_0^T(t) = \mathfrak{F}[R_0^T(t), I_0^T(t), M_0^T(t), W_0^T(t); L_0^T(t), K_0^T(t), H_0^T(t)],$$

is what the mathematicians call a functional.

The results just reached call for numerous observations. Here, I can take up only a few and touch upon them briefly.¹³

First, the reason why I have excluded no element from the categories listed in (6), is that the scholar must never prejudice. Even an economist must first arrive at a complete description of a process and only then decide which elements may be left out because they are economically irrelevant. Nature does not indeed have a cashier's window where we may pay her for the elements (R); yet it would be utterly inept to ignore in the economics of production the fact that natural resources are neither inexhaustible nor uniformly distributed over the globe. The type of economic

model now in vogue, which assumes that growth normally proceeds at a constant rate, simply blots out the most numerous and most critical cases—such as Somaliland or our own Appalachia, for instance. One may feel even more tempted to leave out the waste category, on the ground that waste by definition has no value. But again, as we have come to recognize recently on an increasing scale, the existence of waste is not an innocuous aspect of the economic process.

Second, we should not fail to observe that, since a fund enters a reproducible process and comes out without any impairment of its economic efficiency, service is the only way by which it can participate in the production of the product. While it is true that the cloth—an inflow element—effectively passes into the coat, the same cannot apply to the needle—a fund element. And if one finds the needle in the coat just bought, it certainly is a regrettable accident. The point is that the problem of how the contribution of a fund affects the value of the product is not as directly simple as in the case of a flow factor.

Third, both the value of a fund's service and that of its maintenance flow must, in principle, be imputed to the value of the product. Contrary to Marx's teachings—which have gradually infiltrated the thinking of many a standard economist—there is no economic double counting in this. No worker, no lecturer, can discharge his duties by sending to the shop or to the classroom only that "definite quantity of muscle, nerve, brain, etc., [which] is wasted" during work—as Marx claimed.¹⁴ When one works, one must be present with his entire fund of mental and physical capabilities. The same is true for all other funds: the bridge must be there in its full material existence before we can cross the river. If it were true that we could cross a river on the maintenance flow of a bridge or drive the maintenance flow of an automobile on the maintenance flow of a highway, there would be little financial difficulty in saturating the whole world with all the river crossings and automotive facilities. Economic development could be brought about everywhere with practically no waiting.

IV

As with almost everything else, among the various processes we may envisage in production there is one process that fits the epithet "elementary." It is the process by which every unit of the product—a single piece of furniture or a molecule of gasoline—is produced. The process is

¹³ For greater details see Chapter IX in my forthcoming volume, *Entropy and the Economic Process* (Harvard Univ. Press).

¹⁴ Karl Marx, *Capital* (Chicago, 1932), Vol. I, p. 190.

directly observable in the shop of a cabinetmaker, but it can be easily determined even in a large-scale enterprise. Whatever the product, one thing is certain about the elementary process. In relation to it, most of the funds are idle over large periods of time. The plow is needed only a few days during the whole production time of growing a corn plant; the same is true for the saw or the plane in the production of a table. There is no exception to this rule. And, I contend, one of the most important aspects of the economics of production is how to minimize these periods of fund idleness, whether we are thinking of man, capital equipment, or land.¹⁵

Now, if only one table is demanded during the time of production, T , then obviously we need operate only one elementary process after another in succession. But if n tables, with $n > 1$, are demanded during T , then two alternatives are open to us. We may start n processes at the same time and repeat the operation when they are ended. This is the arrangement in parallel. The second arrangement is the arrangement in line, in which equal batches of processes are begun one after another at intervals equal to an aliquot part of T .

It is obvious that the production function of a system in which the elementary processes are arranged in parallel is obtained from (7) by multiplying every coordinate by n :

$$(8) \quad [n\vec{Q}(t)] = \mathcal{F}\{[n\vec{R}_0(t)], \dots, [n\vec{W}_0(t)]; [n\vec{L}(t)], \dots\}.$$

The point that deserves to be stressed is that the arrangement in parallel offers little or no economic gain. Most kinds of fund factors are now needed in an amount n times as great as in the elementary process. In addition, the idleness of each such fund factor is *ipso facto* amplified by n . The only exceptions are the fund factors that—like a large bread oven, for instance—may accommodate several elementary processes simultaneously. But even though the capacity of such a fund factor would be more fully utilized, its idleness period would remain the same.

The situation completely changes for the arrangement in line. If we assume away any incommensurabilities among the time periods involved in the schedule of the elementary process—an inevitable assumption in practice—and if n is sufficiently large, then a number of processes can be arranged in line so that no fund is idle at any time.¹⁶ The situation is vividly exemplified by an

assembly line, in which every tool and every worker shift without interruption from one elementary process to the next. The arrangement in line, however, describes any factory. In a factory, therefore, the economy of time reaches its maximum. This conclusion opens an avenue of utmost importance. To explore it, we may begin by determining the analytical representation of a factory process.

In a first approach we may consider the entire physical plant as one monolithic fund, P . Over an arbitrary interval $[0, t]$, during which the factory process is in operation, the coordinate of this fund is the function Pt . Similarly, the coordinate of labor power is Ht . And if for the convenience of diction we assume that all flow elements are continuous, their coordinates, too, are represented by linear homogeneous functions. Thus (7) becomes:

$$(9) \quad (\vec{q}t) = \mathcal{G}[(\vec{r}_0t), (\vec{i}_0t), (\vec{m}t), (\vec{w}t); (\vec{P}t), (\vec{H}t)].$$

Let us note that this is a very special functional: first, every function involved in it depends upon a single parameter and, second, the value of t is entirely arbitrary. For these reasons, the functional degenerates into a point function.

There are two degenerate forms. The first is

$$(10) \quad q = \Theta(r, i, m, w; P, H).$$

This formula reminds us of one of the current interpretations mentioned in Section I; namely, that of relation (2). We should note, however, that θ involves two dimensionally different categories of variables. The lower case symbols represent flow rates of some substances. The upper case symbols stand for the rates of service per unit of time. Strangely, however, these last rates do not involve the time element at all: P stands for the plant, and H , for the total labor power—briefly, for quantities of some substances. The second degenerate form is

$$(11) \quad Q = \Psi(R, I, M, W; \mathcal{P}, \mathcal{H}; t).$$

Here, the symbols in roman capital letters are again quantities of some substances; those in script letters are services, and t is the period with which these quantities and services are associated. The form (11), in turn, reminds us of relation (1); i.e., of the quantity interpretation of Wicksteed's formula. The most important difference is that Ψ includes time as an explicit variable.

of the numbers of such processes that can be accommodated at the same time by each unit of the various funds. Batches should be started at intervals of T/d , d being the greatest common divisor of T and of the intervals during which the various funds are needed in an elementary process.

¹⁵ A period of idleness is characterized by the constancy of the corresponding fund coordinate.

¹⁶ The number of elementary processes that should be started each time is the smallest common multiple

This difference bears upon the earlier argument that Wicksteed's production function is homogeneous of the first degree. Actually, Ψ is such a function—as easily follows from the identities

$$(12) \quad Q = qt, \dots, W = wt, \mathcal{P} = Pt, \mathcal{H} = Ht,$$

analogous to (3). There is then an intimate relation between (10) and (11); namely,

$$(13) \quad t\Theta(r, i, m, w; P, H) = \Psi(R, I, M, W; \mathcal{P}, \mathcal{H}; t).$$

Hence,

$$(14) \quad \Theta = [\Psi]_{t=1}.$$

The imbroglia created by (5) is thus resolved. Of course, this does not mean that the factory process operates with constant returns to scale. The homogeneity of Ψ corresponds to the tautology that if we double the time during which a factory works, then the quantity of every flow element and the service of every fund will also double. The issue of returns to scale pertains, instead, to what happens if the fund elements are doubled. The point may be made still clearer.

A superficial inspection of any operating plant suffices to reveal that P consists of some Ricardian land, R , some capital equipment, K , some technical inventories, S , and a special fund, Γ , usually called "goods in process." The last term is definitely a misnomer: half-tanned hides or partly wired radio sets, for example, are not goods. Process-fund seems a more exact term because Γ is in effect a becoming frozen in its various phases. If a photograph of Γ would be projected part by part, as if it were a movie, we would witness the actual change of some input flows into product and waste flows. In spite of this varied composition of any plant, what a given plant can do depends on its blueprint alone, which in turn involves only L and K . And since what a plant can do is shown by the flow rate of its product, we have a first relation

$$(15) \quad q^* = G^*(L, K).$$

A second relation expresses the fact that, given the plant, we require a certain labor power, H^* , if we want to obtain the flow rate q^* . Hence,

$$(16) \quad H^* = H^*(L, K).$$

Should we man the plant with less labor power than H^* , the product flow rate would also become smaller than q^* . To account for this rather common situation, we need to put

$$(17) \quad q = G(L, K, H) \leq q^*.$$

But the fact that this relation looks very familiar should not mislead us: as (17) is defined here, if $q < q^*$, q does not necessarily decrease (and

ordinarily does not) when L and K are decreased while H is kept constant. Actually, the ratio q/q^* measures the percentage of capacity utilized if H is the labor power employed.¹⁷

The next relations are self-explanatory:

$$(18) \quad S = S(L, K, H), \quad \Gamma = \Gamma(L, K, H).$$

There remain the relations binding the other flow elements. The case of the maintenance flow, m , is simple: its size must depend on the amount of equipment to be maintained and the labor fund employed. In addition, by virtue of the conservation law of matter and energy, m must be equal to w_1 , the flow rate of wear-and-tear waste—burned or discarded lubricating oil, broken saw bands, and the like. Hence,

$$(19) \quad m = m(K, H), \quad m = w_1.$$

The same conservation law applies to all other flows. For example, the wood contained in a piece of furniture together with the scrap and the sawdust must exactly account for the wood introduced into the production of that furniture. In the case of a factory system, this relationship yields

$$(20) \quad qt = g(rt, it, w_2t),$$

where w_2 is the flow rate of waste arising from transformation alone. Since (20) must be true for any t , the function g must be homogeneous of the first degree. To this function we may indeed apply the old-time tautology that "doubling the inputs doubles the output." The basic error in some arguments about the returns to scale is to apply this tautology to (17) instead of (20). If L , K , and H are doubled, q does not necessarily double even if we double at the same time all flow inputs. The new factory may be more efficient or more wasteful in using the input materials, which leads us to put

$$(21) \quad w_2 = w_2(L, K, H).$$

Relation (10) is thus decomposed into seven basic relations, listed from (17) to (21), which together constitute the general representation of a factory process.¹⁸

We should now note that the picture at which we have arrived is analogous to the inscription "60 watts" on an electric bulb. That is, it tells us what the factory can do, not what it has done, is doing, or, above all, will do. Like the inscription

¹⁷ As an ordinal measure of the utilized capacity we may use H/H^* . On this point, see note 21, below.

¹⁸ Obviously, this analytical description will have to be completed with additional relations if the particular factory process happens to involve other limitations.

on the bulb, relation (17), for example, is true regardless of whether the factory works or is idle. To show what the factory does, we need an additional coordinate, which, under its various aspects, has deeply preoccupied Marx, but which, perhaps for easily understood reasons, is not found in the analytical tool box of the neoclassical economist. The coordinate is the time, δ , during which the factory works daily. The amount of the daily production, Q , follows immediately from (17):

$$(22) \quad Q = \delta G(L, K, H),$$

a relation which vindicates Marx's dear tenet that labor time measures value even though it has no value itself.¹⁹

V

So much for grounded-in-actuality symbolism. Let me devote my closing remarks to some of the object lessons of this symbolism.

I have stressed the fact that in any elementary process every agent is idle over some definite periods that depend not on our choice or whim but on the state of the arts. I have also argued that we can nonetheless eliminate this kind of idleness completely and that there is only one way to achieve this: to arrange the elementary processes in a factory system. Because of this extraordinary property, the factory system deserves to be placed side by side with money as the two most fateful economic innovations for mankind. I say "economic" and not "technical" because the economy of time achieved by the factory system is independent of technology. Nothing prevents us from using the most primitive technique of cloth weaving in a factory system.

To be sure, there is a second kind of idleness, which depends entirely on our decision: it is the idleness of the factory itself if δ is shorter than a full day. In view of these two kinds of idleness, the economics of production reduces to two commandments: first, produce by the factory system and, second, let the factory operate around the clock.

The first commandment calls for two observations. Even though we can draw the blueprint of a factory for any elementary process whatsoever, not every such factory is economically advantageous. For example, we do not build transoceanic "Queens" by processes in line. The reason is that we can build a "Queen" more quickly than it is demanded in relation to time. The much extolled progress of the industrial revolution may not after all be due only to technological innovations. For these innovations as well as the in-

creased specialization of labor could not have come about unless an increased demand had already induced most craft shops to introduce the system in line. There can be little doubt about it: the factory system was born in an artisan's workshop, not in a factory.

The second observation is that to operate an arrangement of elementary processes in line it is absolutely necessary that we have the freedom to start a process at any time of the day, of the week, and of the year. Unfortunately, we do not always have this freedom. Seasonal variations—which result from the position of our planet relative to our main source of free energy, the sun—prevent us from adapting the factory system to a series of important productive activities. The most important instance is husbandry. For the overwhelming majority of localities, there is a very short and definite period of the year during which a corn plant, for instance, can be grown in the open space from seed. This is why farmers have to work their fields in parallel; that is, in a system of production that yields practically no economy of time. The global analytical representation of that system is (8), not (9). The upshot is that the open-air factories, about which socialist writers in particular have been continuously raving, will remain a utopian dream as long as we are unable to alter the orbit of the earth.

The association between agriculture and the idleness of all agents involved is by now a commonplace. Still, not much is known or even suspected about the importance of the related loss. Two simple illustrations may bring out this point. For the first, let us consider one of the exceptional localities—such as the Island of Bali—where, because of an almost constant climate throughout the year, rice could be grown in an open-air factory. In this case, every day the same number of hands would move over the fields with the same funds of plows, buffaloes, sickles, and flails to plow, sow, harvest, and thresh. Every day the villagers would eat the rice sown that very day, as it were, and they would no longer have to bear the burden of the debts specific to agriculture. But most important of all, we would also discover that, without diminishing the old production at all, there would remain a substantial number of superfluous workers as well as a substantial stock of superfluous equipment—a palpable measure of the overcapitalization of farming in comparison with manufacturing. The second illustration pertains to the current technique by which chickens are raised in the United States. In fact, in this country there are no longer any chicken farms—even though the term continues to be used. Instead, there are chicken factories, with elemen-

¹⁹ Marx, *Capital*, Vol. I, pp. 45, 588.

tary processes arranged in line. The "chicken war" of yesteryear would not have come about if the difference between the old and the new techniques had not been so great as to exceed the shipping cost over the Atlantic plus the wage differential between this country and Europe.

But if not every production activity can be turned into a factory, we should at least try to render the idleness of the agents as small as possible in each particular case. In other words, to bring even a whole economy as near as possible to the functioning of a factory system should be the guiding thought of any planner at any level. In the activity of the countryside, the cottage industry propounded by the agrarians was one answer to this idea. In Romania (so I was told) tractors and drivers shuttle between the plain regions—where two crops are grown each year—and the hilly regions—where only one crop can be grown because of a shorter vegetation period. The necessary funds of tractors and drivers are thus substantially reduced at the cost of some gasoline, oil, and spare parts flows. Less costly solutions would be obtained by mixing several crops within the same locality, the crops being chosen so as to minimize idleness (and hence capital cost). Formally, the problem boils down to splicing graphic patterns with a minimizing condition—a problem of a special type of combinatorial analysis which, I am sure, will prove highly rewarding.

The second commandment is particularly relevant for the underdeveloped economies. In a rich country, it makes perfect sense to operate every factory with one shift, even if the shift be of six or four hours only. In a rich country, there also is no need for night shifts, except whenever technology imposes around-the-clock production. Briefly, in a rich country leisure is a commodity which people may prefer to higher income. Things are different in almost every underdeveloped country where—as every government pronouncement urges—the order of the day is not only development but rapid development. In such countries, the regimen of the eight-hour working day and the reluctance to use night shifts are anachronistic factors that work against the avowed aims.²⁰ There may be many reasons why planners as well as our planning theory have overlooked the

simplest and the most direct lever of economic development; namely, the length of the working day. But one possible reason is that this element of the problem has been left out of the neoclassical representation of a production process. The same omission—we should note—vitiates also the familiar comparisons of the capital-output and capital-labor ratios computed from current statistical data. Since the theoretical apparatus ignores the working time, δ , the most sophisticated statistical agencies, too, have felt no need to include it in their usual collections. Thus, we are unable to obtain valid statistical estimates of K/q^* and K/H^* , the basic technical and theoretical elements.²¹

Another omission of the neoclassical representation is that, as a rule, only the funds (variously defined) are included in the production function. The fact that after a factory is built, production cannot go on unless the input flow factors are forthcoming, has thus been pushed away from the focus of attention. None other than an authority such as A. C. Pigou preached that "in a stationary state factors of production are stocks, unchanging in amount, out of which emerges a continuous flow, also unchanging in amount, of real income."²² The omission of the input flow factors is not unrelated to the present race of all underdeveloped countries to build one factory after another without a thorough examination of the availability of the necessary flow inputs. I am confident that if the prospective economic plan of every country were realized by miracle overnight, we would discover that we have long since been planning for a world with an immense excess capacity of industrial production.

The thoughts I shared with you here may seem simple. Perhaps they are simple, once we have untangled the imbroglia hatched by blind symbolism. The economics of production, its elementary nature notwithstanding, is not a domain where one runs no risk of committing some respectable errors. In fact, the history of every science, including that of economics, teaches us that the elementary is the hotbed of the errors that count most.

²⁰ I may hasten to admit that (22) is only a first approximation formula: a factory working with one shift of ten hours will not produce 25 percent more than with a shift of eight hours. To take better account of facts, we should replace δ by a function of the number of shifts and the number of working hours of each shift. But this amendment does not affect in the least the validity of the statements just made.

²¹ The difficulty is especially serious in the case of comparisons between two different industries. Even if we know that each industry has always used its full capacity, i.e., has worked with the corresponding H^* , the values of capital-labor ratios derived from the usual statistical data are neither comparable nor strictly relevant—unless we also know that both industries employed the same number of shifts. In fact, the Census of Manufactures provides no information on the number of shifts and on the percentage of utilized capacity.

²² Pigou, *op. cit.*, p. 19.

INVESTMENT BEHAVIOR

A MODEL OF NONRESIDENTIAL CONSTRUCTION IN THE UNITED STATES*

By CHARLES W. BISCHOFF
Yale University

I. Introduction

In the rapidly accumulating literature on the determinants of fixed investment in the United States very little ink has been used up in explaining the fluctuations in nonresidential construction. It is not hard to find an explanation for this fact; there is very little available data in which any distinction is made between plant and equipment expenditures. Furthermore, the assumption is often made that plant and equipment are perfect complements and thus can be lumped together (it would also be proper to lump them together if they were perfect substitutes).

Nevertheless, there are several reasons for a separate study of nonresidential construction expenditures. In the postwar period the share of nonresidential construction in total nonresidential fixed investment rose gradually from a low of 32 percent in 1947 to a peak of 40 percent in 1958; since then the share has tended to decline, reaching 30 percent in 1968. An obvious beginning of an explanation might rest on the differential effect of tax policies on the two types of asset. Accelerated depreciation, introduced in 1954, decreased the effective cost of construction more than it decreased the effective cost of equipment. Tax policies undertaken since 1962—the investment tax credit in particular—have favored investment in equipment. In addition, there is reason to believe that monetary policies—and especially the availability of credit—might affect construction more drastically. This factor might be of particular importance in commercial construction, which amounted to more than 26 percent of total nonresidential construction in 1968. Finally, because equipment is often special purpose in nature, and thus a “putty-clay” approach might seem most appropriate for this type of asset, this approach might not be as useful in explaining nonresidential construction; buildings, in general, would seem to be more flexible in terms of alternative uses.¹

* The research described in this paper was carried out under grants from the National Science Foundation and the Ford Foundation.

¹ See [2] for an application of the putty-clay approach to investment in equipment.

As a starting point in this investigation, a relatively simple model of the stock adjustment or accelerator type is considered in Section II. In Section III, relative price effects—including interest rates and accelerated depreciation—are considered. In Section IV the model is modified slightly to allow for the possibility that a putty-clay model applies to the demand for nonresidential construction. Finally, in Section V the effects of capital rationing are considered.

II. A Stock Adjustment Model

Until very recently, national income data on expenditures for nonresidential construction (NRC) have been derived by applying a relatively fixed progress pattern to data on construction contracts. The only published information on such progress patterns is contained in an article by Mayer [8]. Under the assumption that plans are fixed by the time the first significant orders are placed (two months before construction begins, according to Mayer), it seems appropriate to lag all variables by at least this much.

If data were available, it would be preferable to try to explain the basic series on construction contracts (CON) rather than the expenditures series which is derived from it. From Mayer's information on progress patterns it would seem that the following equation holds at least approximately:

$$(II.1) \quad NRC_t = .30 CON_{t-1} + .38 CON_{t-2} + .18 CON_{t-3} + .11 CON_{t-4} + .03 CON_{t-5}$$

Taking these weights as a priori knowledge, an equation to explain NRC should be of the form

$$(II.2) \quad NRC_t = \sum_{i=1}^5 \mu_i X_{t-i}$$

where X_{t-i} is the vector of variables which explain contracts in period $t-i$, and the weights μ_i are given in equation (II.1).

Final decisions to let contracts, made in period t , will depend on the desired stock of buildings in period $t+5$ or $t+6$, when the buildings become available for use. If the desired stock of buildings (KS^*) is a function of expected or planned output,

which itself may be approximated by a linear function of past output levels (Q), we have the relationship

$$(II.3) \quad KS_t^* = \kappa \sum_{j=0}^m \nu_j Q_{t-j}.$$

This desired stock must be compared with the stock of buildings which will be available in period $t+5$ or $t+6$, if no new projects are started. Unfortunately, this stock cannot be measured directly. Instead, I assume that in making the decision to let a contract in period t , the desired stock is compared with the stock of buildings (including the expenditures already made on buildings under construction) at the end of period t , KS_t , defined as:

$$(II.4) \quad KS_t = NRC_t/4 + (1 - \delta/4)KS_{t-1}.$$

This will, in general, be an underestimate of the buildings available in period $t+5$ or $t+6$.

For various reasons, the gap between KS_t^* and KS_t may not be fully closed by contracts let in period t . Assuming that

$$(II.5) \quad CON_t = \lambda(KS_t^* - KS_t) + u_t,$$

where u_t is an error term with properties yet to be specified, we can combine equations (II.2), (II.3), and (II.5) to get:

$$(II.6) \quad NRC_t = \alpha + \lambda \kappa \sum_{i=1}^5 \sum_{j=0}^m \mu_i \nu_j Q_{t-i-j} - \lambda \sum_{i=1}^5 \mu_i KS_{t-i} + u_t.$$

Since the μ weights are assumed to be known, we can define

$$Q_t^W = \sum_{i=1}^5 \mu_i Q_{t-i}$$

$$KS_t^W = \sum_{i=1}^5 \mu_i KS_{t-i}$$

and rewrite (II.6) as

$$(II.6a) \quad NRC_t = \alpha + \lambda \kappa \sum_{j=0}^m \nu_j Q_{t-j}^W - \lambda KS_t^W + u_t.$$

Adding the assumption that u is generated by a first order autoregressive process $u_t = \rho u_{t-1} + \epsilon_t$, with ϵ representing a normally and independently distributed random variable with mean zero and variance σ_ϵ^2 , the specification of the simple stock adjustment model is completed, except for the lag parameters ν_j . Generalizing the polynomial distributed lag method first introduced by Shirley Almon [1] these parameters are assumed to be values of an unrestricted third degree polynomial in j ; that is

$$(II.7) \quad \nu_j = a_0 + a_1 j + a_2 j^2 + a_3 j^3, \quad j = 0, \dots, m.$$

Four weighted averages of past Q 's are introduced into the regression, and the four parameters a_0 , a_1 , a_2 , and a_3 are estimated. For the parameter m , which determines the length of the lag distribution, the values 6, 8, 10, 12, and 14 have been tested.²

Using quarterly data from the national income accounts, for the sample period 1952-68, the results are given in Table 1 (see the Statistical Appendix for details on the data). The results are relatively insensitive to the length of the lag distribution and, judging solely by the goodness of fit, the explanation seems very good. However, there is a great deal of unexplained serial correlation ($\hat{\rho} = .87$ for the best fitting equation). Furthermore, the constant seems quite large, while the speed of adjustment appears to be much too low. In the next section, I introduce additional variables in an attempt to overcome these defects.

III. The Neoclassical Model

The principal competitor to the basic accelerator or stock adjustment model of investment is the neoclassical model, first introduced by Jorgenson [7] seven years ago at these meetings. The firm's demand for capital is derived from some sort of optimizing behavior (maximizing profits or the market value of the firm or minimizing costs), while constrained by some sort of production function. The imputed rental which enters into this marginal analysis is determined by equating the cost of the investment good to the discounted stream of cash flows plus tax benefits generated by the investment good over its lifetime. If the cash flows are assumed to decline exponentially over the life of the investment good, the equality is

$$(III.1) \quad q_i = \int_0^\infty e^{-\tau t} C_{0i}(1 - U)e^{-\delta_i t} dt + q_i U \int_0^{\tau_i} e^{-\tau s} D_i(s) ds,$$

where

- q_i = the price of the i th investment good
- τ = the cost of capital to the firm
- C_{0i} = the initial cash flow generated by the i th investment good
- U = the tax rate on profits
- δ_i = the exponential rate of decline of the cash flow generated by the i th good
- τ_i = the lifetime of the investment good for tax purposes
- $D_i(s)$ = the proportion of the depreciation base

² The estimation has been carried out using the program AUTO, written by Morris Norman and Robert Rasche. The parameter ρ has been estimated using the Cochrane-Orcutt iterative method.

TABLE 1
RESULTS OF ESTIMATING EQUATION (II.6), CAPITAL STOCK ADJUSTMENT MODEL

<i>m</i>	$\hat{\alpha}$	$\hat{\lambda}_K \Sigma \hat{p}$	$\hat{\lambda}$	\bar{R}^2	\overline{SEE}	$\hat{\rho}$	D.W.
6	9.60 (2.45)	.0526 (2.30)	.0807 (1.22)	.9635	.564	.868	2.22
8	9.21 (2.39)	.0487 (2.03)	.0701 (1.04)	.9634	.564	.866	2.20
10	9.00 (2.34)	.0443 (1.73)	.0597 (.85)	.9635	.563	.866	2.20
12	8.96 (2.36)	.0480 (1.73)	.0666 (.90)	.9633	.565	.864	2.18
14	8.95 (2.36)	.0501 (1.63)	.0708 (.90)	.9633	.565	.863	2.18

NOTES: Numbers in parentheses are *t* statistics. The units for \overline{SEE} are billions of 1958 dollars, at annual rates. The statistic \bar{R}^2 is computed as

$$1 - \frac{N-1}{N-K} \left(\frac{\Sigma \hat{e}^2}{\Sigma (NRC_t - \bar{NRC})^2} \right),$$

where *N* is the sample size and *K* is the number of parameters, instead of the more usual

$$1 - \frac{N-1}{N-K} \left(\frac{\Sigma \hat{e}^2}{\Sigma ((NRC_t - \rho NRC_{t-1}) - \bar{NRC}(1 - \rho))^2} \right).$$

for an asset of age *s* which may be deducted from taxable income.

Defining $z = \int_0^T e^{-\tau s} D(s) ds$ as the present discounted value of the depreciation deduction, (III.1) may be solved for C_0 , yielding

$$(III.2) \quad C_0 = \frac{q(1 - Uz)(r + \delta)}{1 - U}$$

Suppose that the production function is homogeneous of degree zero, with a constant elasticity of substitution, σ ; i.e.

$$(III.3) \quad Q = \left(\sum_{i=1}^m \alpha_i X_i \right)^{\sigma/(1-\sigma)},$$

where the *X*'s are factor inputs. Then profit or market value maximization leads to the expression

$$(III.4) \quad X_i^* = \left(\frac{\alpha_i P}{C_{\alpha i}} \right)^{\sigma} Q$$

where *P* = the price of output, for the optimal amount of the *i*th input (X_i^*) to use in producing *Q* units of output. An expression identical to (III.4), except for a proportionality factor, may be derived from a cost minimization model, given the additional assumption that the price of output

is set by applying a constant markup to the minimum average cost of production (see [2] for a derivation).

For a variety of reasons, it is unlikely that an expression such as (III.4) will hold precisely as a demand function for the stock of nonresidential structures. It is not at all clear that an aggregate production function such as (III.3), with constant elasticities of substitution between any pair of factors, can be applied. In addition, there is no allowance for incomplete adjustment due to uncertainty or costs of adjustment. Finally, a substantial proportion of nonresidential construction is carried out by entities which are unlikely to be maximizing profits or market value, or minimizing costs. Public utilities carried out 23 percent of such investment in 1968, while hospitals, religious groups, educational and other institutions accounted for 13 percent more.

For the reasons outlined above, I have preferred to specify simply that

$$KS^* = f(Q, P/C),$$

and to approximate the unspecified functional form with a simple linear function. In addition, I have allowed for separate lag distributions for the output and relative price variables. The model

tested in this section is thus:

$$\text{(III.5)} \quad \text{NRC}_t = \alpha + \sum_{i=0}^m \beta_i Q_{t-i}^W + \sum_{j=0}^n \gamma_j (P/C)_{t-j}^W - \lambda KS_t^W + u_t.$$

As in the previous section,

$$Q_t^W = \sum_{k=1}^5 \mu_k Q_{t-k},$$

$$(P/C)_t^W = \sum_{k=1}^5 \mu_k (P/C)_{t-k},$$

and

$$KS_t^W = \sum_{k=1}^5 \mu_k KS_{t-k}.$$

The error term is again specified to be a first order autoregressive process. As before the lag weights are assumed to be values of an unconstrained third degree polynomial in i or j .

The Statistical Appendix contains details on the series used to construct the relative price term, P/C . Nothing has been said about the choice of the appropriate discount rate. Taking an eclectic approach, I have estimated the parameters of (III.5) using ten different formulations, as follows:

$$\begin{aligned} r_1 &= (RCB)(1 - DU) \\ r_2 &= (2 \times RCB)(1 - DU) \\ r_3 &= (RDP)(1 - DU) \\ r_4 &= (2 \times RDP)(1 - DU) \\ r_5 &= (RCBR)(1 - DU) \\ r_6 &= (2 \times RCBR)(1 - DU) \\ r_7 &= (RDP + GDIV)(1 - DU) \\ r_8 &= (RDP + DGIVR)(1 - DU) \\ r_9 &= .20(1 - U) \\ r_{10} &= .10 \end{aligned}$$

In the listing above, RCB represents the rate of interest on corporate bonds, while $RCBR$ represents the "real" rate of interest on the same bonds (see the Statistical Appendix). RDP is the dividend-price ratio for corporate stocks; in this formulation dividends are taken to represent a measure of the expected earnings of the corporation (under the assumption that corporate managers attempt to set a dividend rate which can be maintained). The variable $GDIV$ adds to the dividend-price ratio a measure of the rate of growth of dividends per share; this formulation is suggested by the fact that if current dividends are at rate R and are expected to grow at rate G , then it can be shown that the stream of future dividends is being discounted at rate $R+G$. The

variable $DGIVR$ is the same except that the rate of growth of dividends is measured in real terms.

All of the first eight cost of capital variables are multiplied by the factor $(1-DU)$ to allow for the deduction of interest payments; the Modigliani-Miller theory [9] suggests that the cost of capital to an unlevered firm should be multiplied by a factor $(1-DU)$, where D represents the desired ratio of debt to total capital in the capital structure of the firm. I have used $D=.2$ as a rough empirical approximation.

The ninth cost of capital is that used by Hall and Jorgenson [4]; it assumes a constant pre-tax cost of capital, with no short-run shifting of the corporate income tax. Finally, the tenth cost of capital assumes constancy of the after-tax cost of capital.

The results of using these different measures, reported in Table 2, show that it makes a good deal of difference which measure is used. Using an F ratio to test the significance of the four weighted variables which represent the lag distribution on relative prices, only the formulations of P/C using costs of capital three, four, nine and ten offer significant improvements over the simple stock-adjustment models. The test statistics are, respectively, 8.59, 10.65, 2.54, and 3.99; they may be compared to the critical value of $F(4, 58)$ which is 2.53 at the 5 percent level. For the best of the alternative models, with twice the dividend-price ratio as the discount rate, Table 3 gives the long-run elasticities of the stock of structures (and investment) with respect to output and the relative price variable at various points in time. The price elasticities would appear to be significantly below unity but above zero. Solving out the lag patterns, Tables 4 and 5 record the adjustment toward the equilibrium stock of structures, given a change in one of its determinants, and also the fluctuations in construction investment as the adjustment is carried out. Compared to other models of this sort the adjustment pattern seems relatively rapid and quite reasonable.

Other aspects of the model seem quite satisfactory as well. The elasticity of capital with respect to output gives a slight indication of increasing returns. The effect of relative prices, while small, is by no means negligible. The fit is quite good and, perhaps most important, the degree of serial correlation of the errors has been reduced quite substantially.

In order to be certain that the success of the relative price indexes based on cost of capital measures 3 and 4 in explaining nonresidential construction is not due solely to the use of stock prices in the denominator of the dividend-price

TABLE 2
RESULTS OF ESTIMATING EQUATION (III.5) WITH VARIOUS
MEASURES OF THE COST OF CAPITAL

Cost of Capital	$\hat{\alpha}$	$\Sigma \hat{\beta}$	$\Sigma \hat{\gamma}$	$\hat{\lambda}$	\bar{R}^2	\overline{SEE}	$\hat{\rho}$	D.W.
1	-24.79 (-1.28)	.0458 (2.02)	2.382 (1.77)	-.0175 (-.213)	.9652	.550	.809	2.30
2	-19.23 (-1.19)	.0411 (1.80)	2.797 (1.79)	-.0237 (-.288)	.9652	.550	.806	2.30
3	-2.62 (-1.66)	.1590 (7.33)	3.665 (5.13)	.4127 (5.79)	.9755	.462	.486	2.22
4	5.03 (5.49)	.1690 (8.93)	4.944 (6.22)	.4560 (7.09)	.9775	.442	.398	2.17
5	-.01 (-.00)	.0376 (.992)	.567 (.479)	.0284 (.212)	.9625	.571	.877	2.29
6	1.35 (.08)	.0359 (.970)	.604 (.562)	.0244 (.193)	.9625	.571	.877	2.28
7	8.62 (2.54)	.0361 (.924)	-.223 (-.285)	.0318 (.282)	.9619	.575	.844	2.24
8	8.62 (2.61)	.0369 (1.000)	-.206 (-.373)	.0327 (.312)	.9623	.573	.839	2.23
9	-45.57 (-1.38)	.1174 (2.83)	11.259 (1.59)	.2064 (2.21)	.9668	.537	.789	2.32
10	-32.64 (-2.03)	.0679 (4.47)	10.758 (2.65)	.1192 (3.09)	.9694	.516	.670	2.29

NOTES: Numbers in parentheses are *t* statistics. The units for \overline{SEE} are billions of 1958 dollars, at annual rates. All results are based on a sample period including the 68 quarters from 1952 through 1968. All of the results are for lag distributions with $m=10$ and $n=10$. For these parameters, which determine the lengths of the two lag distributions, a variety of values between 6 and 14 were tested, but the best results were generally obtained with both parameters equal to 10. The statistic \bar{R}^2 is computed according to the formula given in the notes to Table 1.

ratio, I have added a stock price index to equation (III.5). The variable added is

$$S_t^W = \sum_{k=1}^6 \mu_k S_{t-k}$$

Allowing for lag distributions on S_t^W of six to fourteen quarters, the best results are with a ten quarter lag. Although the coefficients of Q_t^W and $(P/C)_t^W$ are reduced somewhat, these variables retain their statistical significance. In particular, the *F* ratio on the four weighted variables which represent the lagged effects of relative prices is 5.01, compared to a critical value of $F(4, 53)$ of 3.70 at the 1 percent level.

In Table 6, I have used the preferred version of equation (III.5) (with r_t as the cost of capital) to

derive partial effects on nonresidential construction and the stock of structures of the adoption of accelerated depreciation, in 1954. In deriving

TABLE 3
LONG-RUN ELASTICITIES OF STOCK OF STRUCTURES
WITH RESPECT TO ITS DETERMINANTS
(From Equation (III.5), Cost of Capital 4)

Variables Measured as of	Elasticity with Respect to Q	Elasticity with Respect to P/C
1948-IV	.7499	.3245
1958-IV	.7166	.3095
1968-IV	.8304	.2332

TABLE 4

PROPORTION OF ADJUSTMENT OF STOCK OF STRUCTURES
COMPLETED N QUARTERS AFTER SUSTAINED
CHANGE IN Q OR P/C
(From Equation (III.5), Cost of Capital 4)

Quarters after Change	Proportion of Adjustment to Change in Q	Proportion of Adjustment to Change in P/C
0	.000	.000
1	.006	.010
2	.023	.036
3	.050	.075
4	.085	.125
8	.252	.397
12	.497	.718
16	.738	.863
20	.867	.931
40	.996	.998

TABLE 5

FLOW OF INVESTMENT IN CONSTRUCTION N QUARTERS
AFTER SUSTAINED CHANGE IN Q OR P/C ,
AS A PROPORTION OF LONG-RUN FLOW
(From Equation (III.5), Cost of Capital 4)

Quarters after Change	Flow of Investment in Response to Change in Q	Flow of Investment in Response to Change in P/C
0	.00	.00
1	.39	.63
2	1.13	1.74
3	1.81	2.65
4	2.34	3.34
8	3.15	5.65
12	5.21	4.74
16	3.87	2.55
20	2.47	1.77
40	1.05	1.02

TABLE 6

PARTIAL EFFECTS OF ACCELERATED DEPRECIATION ON NONRESIDENTIAL
CONSTRUCTION AND CAPITAL STOCKS

Year	Effects Estimated using Equation (III.5), Cost of Capital 4		Effects Estimated by Hall and Jorgenson	
			in [3]	in [4]
	In Billions of 1958 Dollars	As a Percentage of <i>NRC</i> or <i>KS</i>		
Effects on Nonresidential Construction				
1954	.098	.6%	9.7%	1.9%
1955	.432	2.7	11.4	6.0
1956	.796	4.3	11.1	8.2
1957	.792	4.4	9.9	8.6
1958	.634	3.8	10.2	6.8
1959	.610	3.8	10.4	4.4
1960	.646	3.7	9.7	3.3
1961	.628	3.6	9.7	3.4
1962	.516	3.4	9.4	3.2
1963	.460	2.5	9.3	3.3
1964	.389	2.0		2.9
1965	.364	1.6		2.3
1966	.280	1.2		1.2
1967	— .146	— .6		.6
1968	— .046	— .2		1.4
Total 1954–63	5.612	3.3%	10.1%	4.9%
Total 1954–68	6.453	2.3		3.6
Effects on Stock of Nonresidential Structures (end of year)				
1963	4.275	2.0%	7.2%	3.2%
1968	3.802	1.5		2.5

these numbers, all variables except the present value of the depreciation deduction have been held at their actual levels. Thus, any feedback on construction via the multiplier or any other mechanism is ignored. The overall effect of accelerated depreciation might be either larger or smaller than these derived partial effects (it might be smaller if, for example, monetary policy would have been less restrictive in the absence of accelerated depreciation). Nevertheless, the partial effects are of some interest, if only in making comparisons to the partial effects derived from other models using similar assumptions. In Table 6 a comparison is made to the partial effects derived in two papers by Hall and Jorgenson [3] [4]. The comparison shows that both of the Hall-Jorgenson papers indicate larger effects, but that the discrepancy is smaller when the results I have derived are compared to their most recent paper [4].

IV. A Putty-Clay Model

A different line of argument against naive application of a model such as the one represented by equation (III.5) is suggested by capital-theoretic models in which factor proportions are variable only before fixed capital goods are put into place. These models (putty-clay models) can be shown, at least given a few additional assumptions, to imply that the dynamic effects on investment of a change in output will differ considerably from the dynamic effects of a change in relative prices (see [2]). In a model of this sort output changes will result in the familiar acceleration effect, in which investment rises or falls sharply in response to the necessity to increase or decrease capacity; then tapers off to the level required to sustain the new level of capacity. No such acceleration effect, however, would be visible in response to a change in relative prices.

As reported in Tables 4 and 5, the response to a change in P/C is very similar to the response to changes in Q ; strong acceleration effects appear for both variables. This would suggest rejection of the putty-clay hypothesis. This result, however, may be an illusory one built into the model. Although the lag weights β_i and γ_i in equation (III.5), which affect the lag distributions for the first 15 periods, are indeed permitted to differ, the introduction of KS_t^W into the equation requires that the lag weights for both output and relative prices must decline exponentially after the 15th period. This restriction means that the putty-clay hypothesis cannot really be tested.

In order to allow for more flexibility in the lag distributions, I have tested the effects of adding KS_{t-1}^W , NRC_{t-1} , NRC_{t-2} , and various combina-

tions of these variables to equation (III.5). The addition of these variables allows for different lag distributions for output and relative prices in the long run as well as the short run. The results of this experiment are negative; in all cases the acceleration effect for changes in relative prices remains. Thus, the putty-clay hypothesis apparently does not apply to investment in non-residential structures.

V. The Effects of Capital Rationing

In recent doctoral dissertations by Hand [5] and Jaffee [6] an index of capital rationing has been derived. The index is based on several ratios of the proportion of small business loans to total loans and the proportion of loans granted at the prime rate. It is available starting in 1952. This index is a pure number; in order to give it a dollar dimension it is multiplied by the current dollar stock of outstanding commercial loans and deflated by the price index for NRC .

This scaled index has been added to various versions of equation (III.5), with a sample period starting in 1954. In most cases the coefficient of the capital rationing index has the correct negative sign, but the coefficients are very small and in no case do they approach statistical significance. I conclude that, at least using this index, there is no clear evidence of an effect of capital rationing on nonresidential construction.

VI. Conclusions

In this paper I have developed a model for the explanation of expenditures for nonresidential construction based on a linear approximation to a neoclassical model similar to Jorgenson's. The model does a good job of explaining the fluctuations in this series. Although the long-run elasticity of construction expenditures with respect to relative prices is much less than unity, the effects are statistically significant and sizable in the short run. The dividend-price ratio appears to be the best measure of the cost of capital for use with this series. Attempts to revise the model in accord with the putty-clay hypothesis and to allow for the effects of capital rationing have led to negative results.

STATISTICAL APPENDIX

The variables Q , P , NRC , and q are taken directly from the National Income and Product Accounts. Seasonally adjusted data are used, and flow variables are in billions of 1958 dollars, at annual rates. Q is Business Gross Product and P is the deflator for that series. NRC is Investment in Nonresidential Structures and q is the deflator

for that series. The series KS is constructed from $P\text{NRC}$ via the equation

$$KS_t = \text{NRC}_t/4 + (1 - \delta/4)KS_{t-1}$$

where δ is the yearly rate of depreciation. Bench marks for KS at the end of 1946 (118.7) and the end of 1964 (218.6) are taken from the OBE capital goods study, and the rate of depreciation (.0608) is derived so as to make the stock equal those two bench marks at the appropriate dates.

In constructing the quasi-rent on structures, defined in equation (III.2), the variable U is defined as the statutory marginal tax rate on profits. The present value of the depreciation deduction, z , is defined according to

$$z_t = (1 - w_t)PVSL + w_tPVSVD$$

where $PVSL$ is a continuous approximation to the present value of the depreciation deduction using straight line depreciation,

$$PVSL = \int_0^{\tau} \frac{1}{\tau} e^{-rs} ds,$$

and $PVSVD$ is a continuous approximation to the present value of the depreciation deduction using the sum-of-the-years'-digits method,

$$PVSVD = \int_0^{\tau} \frac{2(\tau - s)}{\tau^2} e^{-rs} ds.$$

$PVSVD$ is used as an approximation to the present value using either SYD or the double declining balance method, since the present values are very close together. Accelerated amortization on emergency structures constructed during the Korean war is ignored. The variable w represents the proportion of depreciation taken by accelerated methods, and is based on the "learning curve" analysis of Wales [10]; starting in 1954

$$w_t = (1 - .79^{t'})$$

where the time index t' is measured in years and equals 0 at the middle of 1952. The lifetime of structures for tax purposes, τ , is taken to be 27.8 years through 1954 declining to 22.8 years by the end of 1956, following Hall and Jorgenson [3].

The discount rates are described in the text. The "real" interest rate on bonds ($RCBR$) is derived by subtracting .015 from the bond yield

until 1965; from then on

$$(6p_t + 5p_{t-1} + 4p_{t-2} + 3p_{t-3} + 2p_{t-4} + 1p_{t-5})/21$$

is subtracted from the bond yield. The rationale is that inflationary expectations were not widely held until the recent inflation, which accelerated around 1965.

The stock price index used in Section III is Standard and Poor's Index of 500 Stocks. The capital rationing index and the stock of outstanding commercial loans were supplied by Professor Dwight Jaffee, of Princeton University.

REFERENCES

1. SHIRLEY ALMON, "The Distributed Lag between Capital Appropriations and Expenditures," *Econometrica*, Jan., 1965, pp. 178-95.
2. CHARLES W. BISCHOFF, "Lags in Fiscal and Monetary Impacts on Investment in Producers' Durable Equipment," Cowles Foundation Discussion Paper No. 250 (New Haven, Conn.), June 11, 1968 (to be published by the Brookings Institution in *Tax Incentives and Capital Spending*).
3. ROBERT E. HALL AND DALE W. JORGENSON, "Tax Policy and Investment Behavior," *A.E.R.*, June, 1967, pp. 391-414.
4. ———, "The Quantitative Impact of Tax Policy on Investment Expenditures," Working Paper 136, Institute of Business and Economic Research, Univ. of California, Berkeley, July, 1968 (to be published by the Brookings Institution in *Tax Incentives and Capital Spending*).
5. JOHN HAND, "The Availability of Credit and Corporate Investment," unpublished dissertation, Massachusetts Institute of Technology.
6. DWIGHT JAFFEE, "Credit Rationing and the Commercial Loan Market," unpublished dissertation, Massachusetts Institute of Technology.
7. DALE W. JORGENSON, "Capital Theory and Investment Behavior," *A.E.R.*, May, 1963, pp. 247-59.
8. THOMAS MAYER, "The Inflexibility of Monetary Policy," *Rev. of Econ. and Statis.*, Nov., 1958, pp. 358-74.
9. MERTON H. MILLER AND FRANCO MODIGLIANI, "Some Estimates of the Cost of Capital to the Electric Utility Industry, 1954-57," *A.E.R.*, June, 1966, pp. 333-91.
10. TERENCE J. WALES, "Estimation of an Accelerated Depreciation Learning Function," *J. of the A.S.A.*, Dec., 1966, pp. 995-1009.

TAX POLICY AND INVESTMENT EXPENDITURES IN A MODEL OF GENERAL EQUILIBRIUM*

By LAURITS R. CHRISTENSEN

University of Wisconsin

Introduction

It has long been an article of faith among economists that investment incentives such as tax credits and liberal depreciation rules are effective in stimulating investment expenditures. Recently the faithful have been rewarded with the emergence of a body of empirical evidence purporting to support their beliefs. The seminal paper by Hall and Jorgenson [7] has caused a revival of interest in assessing the effectiveness of tax policy in stimulating investment. In 1967 the Brookings Institution sponsored the Conference on Effect of Investment Tax Credit and Accelerated Depreciation. In addition to further results by Hall and Jorgenson [9] there were contributions by Bischoff [1], Coen [3], and Klein and Taubman [11]. The papers differed somewhat in technique and emphasis, but the conclusions of the authors unanimously supported the thesis that tax policy can be and has been successful in stimulating investment expenditures in the U.S. In his discussion of these papers Harberger [10] questioned how this conclusion could be reconciled with the fact that the share of investment in GNP has remained unperturbed by the tax stimuli. He suggested that the answer lay in macroeconomic constraints, in particular the saving-investment identity, which did not appear in the partial equilibrium analyses of the authors. Harberger concluded that tax policy had not been effective in stimulating gross private domestic investment.

Taubman and Wales [13] attempted "to analyze in a general equilibrium context the long-run steady state implications of investment subsidies in general, and of the tax credit and of a change in depreciation methods in particular." In their model, tax policy could not have an effect unless consumption and saving had nonzero interest elasticities or the propensities to consume wage and property income differed. Their analysis of tax policy proceeded under the assumptions that saving was proportional to disposable income and that the proportionality factor was an increasing function of the rate of interest. They concluded

that investment stimuli were effective but less than a similar partial equilibrium analysis would indicate.

In this paper the questions raised by Harberger and Taubman-Wales are pursued in the context of a simple macroeconomic model of the U. S. economy. In the studies to date, attention has centered on a detailed specification of the production sector. The household and government sectors have been vaguely specified, if at all, and the integration of all sectors by social accounting relations has not been described. Saving and investment must be equal *ex post*; thus it is important to specify the determination of saving as well as investment.

The Macroeconomic Model

Recent contributions to the literature on the effectiveness of investment incentives, e.g., [2] [5] [8], have been primarily concerned with the role of the elasticity of substitution between labor and capital and the pattern of lagged adjustment of capital stock to output and price changes. These controversies are avoided by assuming that there is an aggregate production function exhibiting constant returns to scale and assuming that the discrete time unit is long enough for adjustments to be fully worked out.¹ Assuming that labor and capital are paid their marginal products and that there is competitive pricing of output, revenue to the producing sector is exhausted in payments to capital and labor. Profits, appropriately defined, and the present value of the production sector are zero. All firms exhibit constant returns to scale, and hence the size of individual firms is indeterminate. Firm managers take the capital stock, wages, and prices as given and optimize with respect to output and labor input. Firm managers have no investment demand function; they accept whatever additions are made to capital stock by firm owners. Firms can produce two types of output as joint products, consumption goods, and investment goods. The output

* The author would like to thank S. S. Christensen, A. S. Goldberger, and D. W. Jorgenson for helpful comments. Research support has been provided by the National Science Foundation and the Graduate School of the University of Wisconsin.

¹ Lucas [12] has developed a model in which (a) holds but (b) does not. He cited several studies from the industrial organization literature which support the constant returns hypothesis. The absence of an investment function, as discussed below, obviates the need to consider firms' adjustments to "desired" capital stock.

mix is part of the manager's optimizing decision. Under these conditions there are consumption and investment goods supply functions and a labor demand function. The arguments of the functions are: the capital stock K ; the prices of the four outputs and inputs, consumption goods Q_{CS} , investment goods Q_{IS} , capital services P_{KS} , and labor services P_{LD} ; and the technology or state of the arts index A . The model presented here explains only relative prices; thus one of the four prices, say Q_{IS} , can be thought of as predetermined. It may be convenient to think of Q_{IS} or the price level as determined by monetary policy, which has no explicit role in the model. To simplify the exposition, assume that the aggregate flow of capital services is proportional to the aggregate capital stock at the end of the previous period. The symbol K_{t-1} is used to represent both quantities, with units of measurement appropriate to the context. The three demand and supply functions and the zero profit condition can be written

$$(1) \quad C_S = C_S(K_{t-1}, Q_{CS}, Q_{IS}, P_{KS}, P_{LD}, A)$$

$$(2) \quad I_S = I_S(K_{t-1}, Q_{CS}, Q_{IS}, P_{KS}, P_{LD}, A)$$

$$(3) \quad L_D = L_D(K_{t-1}, Q_{CS}, Q_{IS}, P_{KS}, P_{LD}, A)$$

$$(4) \quad Q_{CS}C_S + Q_{IS}I_S = P_{KS}K_{t-1} + P_{LD}L_D.$$

A , Q_{IS} , and K_{t-1} are predetermined; the remainder of the variables in (1)–(4) are endogenous.

The production sector behavioral relations interact with the household sector to determine the general equilibrium levels of the outputs, labor input, and relative prices. In Christensen [4] a utility maximization approach has been developed which yields a consumption demand function and a leisure demand function. Expenditures on consumption and leisure are proportional to total resources; the proportionality factors are functions of the prices of consumption and leisure and the rate of return on household wealth. The present discussion is limited to the case in which the proportionality factors are independent of the rate of return. The concept of total resources to be allocated between the present and the future is nonhuman wealth (at the end of previous period) W_{t-1} plus potential human wealth, which consists of all potential working time evaluated at the expected after-tax wage rate discounted to the present. The simple expectational assumption that after-tax wage rate and population will grow at constant rates allows the use of current period value as a proxy for the total. Thus the total resource measure is $W_{t-1} + \Omega P_{LS}H$, where P_{LS} is the after-tax wage rate, H is time available, and Ω is a function of the expected rates of growth of P_{LS} and H and the discount rate. The consumption

demand function can be written

$$(5) \quad Q_{CD}C_D = C_D(Q_{CD}, P_{LS}) \cdot (W_{t-1} + \Omega P_{LS}H),$$

where Q_{CD} and C_D are the price and quantity of household sector consumption; and the leisure demand function can be written

$$(6) \quad P_{LS}J = J(Q_{CD}, P_{LS}) \cdot (W_{t-1} + P_{LS}H),$$

where J is the amount of leisure consumed by the household sector.

The labor supply function of the household sector L_S is implicit in the leisure demand function and the time constraint

$$(7) \quad H = L_S + J.$$

Equations (1)–(7) are the heart of the macro-model. They contain all the behavioral relations which are specified. The remaining equations serve to tie the sectors together and to make the model dynamic.

The specification of the household or consumer sector is completed with a saving function. Saving in period t is defined to be the change in wealth less accrued capital gains on wealth. Private national wealth is composed of the value of tangible assets $Q_{IPD}K$ and net claims on governments and foreigners $Q_D D$,

$$(8) \quad W = Q_{IPD}K + Q_D D.$$

The saving relation is

$$S_t = W_t - W_{t-1} - K_{t-1}\Delta Q_{IPD} - D_{t-1}\Delta Q_D.$$

Denoting gross private domestic investment as I_{PD} and net foreign investment plus the government deficit as F and assuming that a constant proportion μ of K must be replaced annually provides two recursive relationships:

$$(9) \quad K_t = I_{PD} + (1 - \mu)K_{t-1}$$

$$(10) \quad D_t = F_t + D_{t-1}.$$

Substituting W , K , and D into the saving relation yields

$$(11) \quad S_t = Q_{IPD}I_{PD} - \mu Q_{IPD}K_{t-1} + Q_D F_t.$$

Saving is equal to gross private domestic investment less replacement plus net foreign investment plus the government deficit.

Since interrelations among the production, household, and government sectors play a major role in the ensuing discussion, it is convenient to list the identities which tie the three sectors into a complete model:²

² Henceforth the term government is used to represent the government and foreign sectors. Behavior in both sectors is exogenous, but the primary interest here is in the relation of the private domestic sector to the government sector. Thus for simplicity the foreign sec-

$$(12) \quad I_S = I_D + I_G + I_V.$$

The output of durable goods and structures is equal to sales to the production sector I_D and the government sector I_G , plus the increase in inventories I_V .

$$(13) \quad C_S = C_D + C_G + C_V.$$

The output of nondurable goods and services is equal to sales to the household sector C_D and the business sector C_G , plus the increase in inventories C_V . Following the convention in the U. S. national accounts gross private domestic investment I_{PD} is defined as

$$(14) \quad P_{PD} = I_D + I_V + C_V.$$

There are value analogs to all three of these equations

$$(15) \quad Q_{IS}I_S = Q_{ID}I_D + Q_{IG}I_G + Q_{IV}I_V$$

$$(16) \quad Q_{CS}C_S = Q_{CD}C_D + Q_{CG}C_G + Q_{CV}C_V$$

$$(17) \quad Q_{IPD}I_{PD} = Q_{ID}I_D + Q_{IV}I_V + Q_{CV}C_V.$$

Labor services are supplied by the household sector to the production and government sectors

$$(18) \quad L_S = L_D + L_G$$

$$(19) \quad P_{LS}L_S = P_{LD}L_D + P_{LG}L_G.$$

Equations (1)–(19) form a complete model determining the nineteen endogenous variables

- | | | | |
|---------------|--------------|--------------|--------------|
| 1) C_S | 2) I_S | 3) L_D | 4) P_{KS} |
| 5) Q_{CD} | 6) P_{LS} | 7) J | 8) W |
| 9) K | 10) D | 11) S | 12) I_D |
| 13) C_D | 14) I_{PD} | 15) Q_{ID} | 16) Q_{CS} |
| 17) Q_{IPD} | 18) L_S | 19) P_{LD} | |

Saving, Investment, and Investment Incentives

As in previous studies, investment incentives enter the formulation via the relation between capital asset prices and capital service prices. For a homogeneous capital good, which depreciates exponentially and is unaffected by taxes, the relation between the asset price q and the service price p can be specified

$$q_t = \sum_{m=t}^{\infty} \prod_{s=t+1}^{m+1} \frac{1}{1+r_s} p_{m+1} (1-\mu)^{m-t}$$

where μ is the rate of replacement and r_s is the rate of return at time s . The quantity of capital services at time $m+1$ from one unit of investment at time t is $(1-\mu)^{m-t}$. The sequence of capital services declines geometrically, 1, $(1-\mu)$,

tor is suppressed by combining it with the government sector. In order to highlight the structure of the model index number problems and aggregation adjustments are ignored in this presentation.

$(1-\mu)^2 \dots$. The equation in p and q can be rewritten in the recursive form

$$q_t = \frac{1}{1+r_{t+1}} [p_{t+1} + (1-\mu)q_{t+1}],$$

which can be inverted to obtain

$$(*) \quad p_t = q_{t-1}r_t^* + q_t\mu,$$

where

$$r_t^* = r_t - \frac{q_t - q_{t-1}}{q_{t-1}}$$

is the own-rate of return on capital. Thus the service price is composed of the real cost of capital plus the current cost of replacement. Given the sequence of asset prices, the rate of return, and the rate of replacement, the service price can be computed. Alternatively, given the sequence of asset prices, the service price, and the rate of replacement, the implicit rate of return can be computed.

Several studies have been made in which equations such as (*) have been used to relate aggregate variables. Usually tax parameters are introduced which apply corporate tax regulations to the whole economy. A more desirable approach would be one which permits a detailed specification of tax parameters but allows easy integration with an aggregate model such as the one outlined above—in order to compute general equilibrium results from changes in tax policy.

Christensen and Jorgenson [5] have developed a detailed accounting system for capital input. They distinguish seventeen different capital goods by sector and type. Each capital service price can be written

$$p_i = q_i e_i$$

where q_i is the corresponding asset price and e_i is a function of the cost of capital, the rate of replacement, and the tax treatment for the i th capital good. Define the aggregate price of capital services and price of assets in terms of the detailed types as follows

$$P_{KS} = \prod p_i^{w_i}, \quad Q_{IPD} = \prod q_i^{v_i},$$

where the w_i are service flow value shares and the v_i are investment value shares. This yields the relation

$$(20) \quad P_{KS} = \prod (q_i e_i)^{w_i} = Q_{IPD} X E,$$

where $X = \prod q_i^{w_i - v_i}$ is the aggregation error due to differences in service and investment weights, and $E = \prod e_i^{w_i}$ is a geometric mean of the e_i .

Equation (20) can be added to the macromodel; the additional endogenous variable is the rate of

return on capital after corporate taxation but before personal taxation, which is implicit in the complicated expression E . In essence equation (20) allocates the before tax rate of return between the government (tax) share and the net return to capital. Tax rates on corporate income and property income, allowable depreciation formulas and asset lifetimes, and investment tax credits all influence this allocation. As a first approximation one can think of the before tax return as determined simultaneously with P_{KS} and Q_{IPD} in the first nineteen equations; then the net return is determined in (20) by the tax parameters which reflect the current tax policy. This would suffice as a year-to-year description of the solution to the model only if the solution to equations (1)–(19) were independent of changes in tax policy. Under such conditions the inescapable conclusion would be that tax policy has no impact on investment expenditures. Thus it is important to analyze the feedback effects of tax policy in the full model.

A full analysis of the impact of tax policy on investment (and the other endogenous variables) would require a set of simulation results on the model outlined above. Such an investigation is indeed in progress; however, it may be instructive to make a theoretical analysis of the model before allowing the numbers to tell their own story. In this spirit let us consider two cases which are relatively easy to trace through the model and which lead to radically different conclusions regarding the effectiveness of investment incentives.

There are myriad fiscal packages which include investment incentives. Taubman and Wales [13] have used a budget-balancing package in which revenue lost by the investment incentives is made up by a proportional income tax increase. There are many other ways that the budget could be balanced, but budget-balancing packages are certainly not the only ones of interest. In fact the case of no offsetting revenue or expenditure measures is of central interest because several previous studies have not specified explicitly any government action in addition to the investment incentives.

First, consider the case where the foregone revenue of the investment incentives results in a corresponding deficit in the government budget and that the deficit is covered by the sale of government debt to the household sector. Assume that the monetary authorities react in order to keep the money supply unchanged.

The initial impact of this fiscal policy is limited to equations (8), (10), (11), and (20). The impact on (20) is the tendency for the after-tax rate of

return to rise; while the impact on (8), (10), and (11) is the tendency for the government deficit to rise, which in turn causes saving and end-of-period wealth to rise. In fact it can be seen that in the first period the after-tax rate of return will rise to completely offset the effects of investment incentives on (20), leaving the relation between P_{KS} and Q_{IPD} unchanged. The behavioral equations for the production and household sectors are therefore unaltered. Gross private domestic investment is unaltered because the entire increase in saving is "invested" in government debt. Moving to the second time period there is a direct effect on the household sector behavioral equations; wealth has increased, thus both consumption and leisure tend to increase proportionately. In the second period and subsequent periods there will be feedback effects throughout the model, but the impact should not be reversed. Consumption and leisure increase, both causing saving and investment to decline. A further effect which is not made explicit in the model is envisaged. The increase in the after-tax rate of return on tangible capital causes the household sector portfolio to be out of balance. Tangible capital is now more attractive relative to government debt. The rate of interest on government debt must rise to insure that the full amount of the debt be willingly held. With a higher rate of interest the household portfolio can regain balance. This mechanism need not be explicit because the interest rate has no feedback effects in the model. The conclusion for this first case, which may seem startling, is that the imposition of investment incentives causes gross private domestic investment to be smaller than it otherwise would have been.

Second, consider the case where the foregone revenue of investment incentives is balanced by a corresponding reduction in government expenditures. In particular let the expenditure reduction be composed entirely of durable goods and structures. Again there is the impact on (20) in which the after-tax rate of return tends to rise. There is no impact on relative prices, so the behavioral equations of the production sector are unaffected. Specifically, the output mix of consumption and investment goods is unaltered. But now the amount of investment goods purchased by the government sector has declined, leaving a larger amount for gross private domestic investment. The increase in the rate of return makes tangible capital more attractive and the investment incentives, which amount to a tax cut, provide the funds to increase the household sector's holdings of tangible capital. After tax property income increases by exactly the amount needed

to purchase the investment goods no longer "claimed" by the government. Thus the impact of the investment incentives is a dollar-for-dollar increase in gross private domestic investment; each dollar of tax revenue lost through the investment incentives results in a dollar increase in gross private domestic investment. In subsequent periods such a policy causes faster rates of growth of tangible capital and private wealth due to increases in investment. Consequently income and consumption can also grow faster. The conclusion for this second case is that investment incentives accompanied by reductions in government expenditures for durable goods and structures provide a powerful tool for stimulating gross private domestic investment.

The extreme disparity of results for the two cases discussed above suggests that a complete specification of government policy is a necessary prerequisite to drawing any conclusions regarding the effectiveness of a single policy tool. Previous studies suffer from the lack of such specifications. The model presented here is not directly comparable with previous work because it lacks investment demand functions by firms. The model indicates, however, that studies which claim to demonstrate the effectiveness of investment incentives probably entail strong assumptions about the (unspecified) supply side of the market for capital goods.

Summary and Conclusion

A complete macroeconomic model has been specified in order to investigate the effects of investment incentives in a general equilibrium setting. Constant returns to scale are assumed for the production sector implying the lack of an investment demand function. Investment is determined through firms' decisions to produce investment goods and households' saving and portfolio decisions. The saving decision results from an intertemporal utility maximization. The effectiveness of tax policy in stimulating private investment is found to depend critically on the form of the complete fiscal package. At one extreme investment incentives have no immediate impact and actually have a lagged effect causing investment to decline. This occurs if the resulting increase in saving is entirely "invested" in government debt. Consumption increases due to higher wealth, and investment suffers. At the other extreme investment incentives initially increase

private investment dollar for dollar. This occurs if the government revenue foregone via investment incentives is matched by a concurrent reduction in government purchases of investment goods from the production sector.

The conclusion of this paper is that judging the effectiveness of tax policy requires much more attention to the development of an appropriate macroeconomic context and the consideration of complete fiscal packages than has been evidenced in previous studies. Some theoretical outcomes of tax policies have been suggested. Empirical implementation of the model outlined here to measure the effects of various tax policies requires much additional research.

REFERENCES

1. C. W. BISCHOFF, "Elasticities of Substitution, Capital Malleability and Distributed Lag Investment Function," in a forthcoming volume by the Brookings Institution.
2. R. M. COEN, "Tax Policy and Investment Behavior: Comment," *A.E.R.*, June, 1969, pp. 370-79.
3. ———, "Tax Policy and Investment in Manufacturing, 1954-66," in a forthcoming volume by the Brookings Institution.
4. L. R. CHRISTENSEN, "Saving and the Rate of Return" (Ph.D. dissertation, Univ. of California, Berkeley, 1968).
5. L. R. CHRISTENSEN AND D. W. JORGENSEN, "The Measurement of U.S. Real Capital Input, 1929-67," *Rev. of Income and Wealth*, Dec., 1969.
6. R. EISNER, "Tax Policy and Investment Behavior: Comment," *A.E.R.*, June, 1969, pp. 379-88.
7. R. E. HALL AND D. W. JORGENSEN, "Tax Policy and Investment Behavior," *A.E.R.*, June, 1967, pp. 391-414.
8. ———, "Tax Policy and Investment Behavior: Reply and Further Results," *A.E.R.*, June, 1969, pp. 388-401.
9. ———, "The Quantitative Impact of Tax Policy on Investment Expenditures," in a forthcoming volume by the Brookings Institution.
10. A. C. HARBERGER, "Tax Stimuli and Investment Behavior," in a forthcoming volume by the Brookings Institution.
11. L. R. KLEIN AND P. TAUBMAN, "Impact of Accelerated Depreciation and Investment Tax Credit on Investment and General Economic Activity," in a forthcoming volume by the Brookings Institution.
12. R. E. LUCAS, "Adjustment Costs and the Theory of Supply," *J.P.E.*, Aug., 1967, pp. 321-34.
13. P. TAUBMAN AND T. J. WALES, "Impact of Investment Subsidies in a Neoclassical Growth Model," *Rev. of Econ. and Statis.*, Aug., 1969, pp. 287-97.

CAPACITY, OVERTIME, AND EMPIRICAL PRODUCTION FUNCTIONS*

By ROBERT E. LUCAS, JR.

Carnegie-Mellon University

I. Introduction

There are at present two leading theories describing the response of a competitive industry to cyclical fluctuations in product demand, one based on the neoclassical production function, the other on the hypothesis of fixed factor proportions. In this introductory section, the implications of each for cyclical movements in output, inputs, and relative prices are sketched, and the formidable (and largely familiar) empirical cases against both theories are reviewed.¹ In the remainder of the paper, an alternative theory, consistent with the evidence which appears to contradict the first two theories, is articulated.

To review the two theories in their simplest form, consider a competitive industry with many firms producing a single output by means of two inputs: labor and capital. Production is subject to constant returns to scale:

$$(1) \quad y_t = f(x_t),$$

where y_t is output per unit of capital in period t , and x_t is labor (man-hours) per unit of capital. In the current period, capital is fixed at its beginning of period level, so that the current period, short-run decision problem for each firm is: choose labor and output, subject to (1), so as to maximize profit per unit of capital,

$$(2) \quad p y_t - w x_t.$$

The wage rate w_t and the output price p_t are regarded as parameters by firms. When a particular production function is specified, solution of this maximum problem yields theories of short-run output supply and labor demand.

If the production function is neoclassical ($f'(x) > 0$, $f''(x) < 0$ for all $x > 0$) one obtains the following well-known implications: (a) the marginal cost curve is smooth and upward sloping, so that demand shifts induce output and price changes in the same direction; (b) real wages $w/p = f'(x)$, should move countercyclically (be negatively correlated with x); (c) empirical esti-

mates of (1) should yield an output-labor elasticity $x f'(x)/f(x)$ roughly equal to labor's share.

An alternative theory is obtained from maximizing the quantity (2) when f exhibits fixed factor proportions:²

$$(3) \quad y_t = \min[x_t, 1].$$

The principal implications of this theory are: (a) the marginal cost curve is flat, up to capacity output, so that demand increases will lead first to expansions of output, next to price increases; (b) since capital earns a return only if $x_t = 1$, this equality should hold at cyclical peaks; (c) both labor's share and the real wage should move countercyclically; (d) under an optimal policy, $y_t = x_t$, so that empirical estimates of (1) should yield an output-labor input elasticity of unity.

Evidence from direct estimation of production functions clearly favors the fixed proportions theory. The table below reports six estimates of the Cobb-Douglas form of (1), with an exponential trend term added.³ In each case, the estimated output labor input elasticity differs insignificantly from unity. If the test for a "good" production function is an elasticity near labor's share, these results must be discarded (as one suspects similar results have been by many others). Yet by usual econometric standards, these results are excellent, and they exhibit a satisfying uniformity over a wide variety of samples and time periods. Perhaps it would be more interesting to search for a theory which can account for this regularity.

Evidence on cyclical movements in real wages

* Equation (3) assumes that capital and labor units have been chosen so that peak output per unit of either input is unity.

² The results in Table 1, lines 1-3 are based on annual U.S. time series from Kendrick [5]. Line 4 is based on annual data from OBE sources. Lines 5 and 6 use OBE quarterly data. They differ only in the capital variable, line 5 using a one-hoss-shay (OHS) assumption to cumulate investment, and line 6 using declining balance (DB). The series are described in detail in [7], an earlier draft of this paper. (They may be obtained from the author on request.) The estimates are obtained using the two-step method proposed by Durbin in [3], which is appropriate under the hypothesis of first-order autocorrelation in the residuals. The statistic R^2 is one minus the sum of squared errors from the second stage, divided by the sum of squared deviations of the dependent variable from its mean (not the computed R^2 from the second stage).

* I wish to thank my colleagues L. A. Rapping, Myron Joseph, T. W. McGuire, and Wynn Winkler for their comments and suggestions. I would also like to thank Arvind Jain for his assistance and the National Science Foundation for its support.

¹ The review was suggested by, and resembles, some of Phelps's observations in [9].

TABLE 1
ESTIMATED PRODUCTION FUNCTIONS

Sample	Period	Constant	$\frac{L}{K}$ -elasticity	Trend	$\hat{\rho}$	DWS	R ²
Kendrick...	1891-1953	-.221 (.017)	1.141 (.079)	.0058 (.0003)	.730 (.094)	2.23	.978
Kendrick...	1891-1929	-.395 (.047)	1.035 (.167)	.0084 (.0011)	.533 (.145)	2.23	.793
Kendrick...	1931-1953	-2.909 (.710)	.973 (.060)	.0193 (.0009)	.345 (.200)	1.78	.993
OBE.....	1930-1965	-.314 (.018)	1.039 (.027)	.0123 (.0003)	.571 (.110)	1.82	.967
Quarterly... (OHS)...	1948-II-1965-IV	.420 (.046)	.992 (.041)	.0027 (.0002)	.579 (.115)	2.13	.996
Quarterly... (DB)....	1948-II-1965-IV	.402 (.049)	.974 (.048)	.0026 (.0002)	.580 (.115)	2.10	.990

is inconsistent with both theories. Thus Bodkin concludes his recent review and extension of the evidence on this point with: "... the majority of the analyses performed with the U.S. data support the ... view that real wages are positively related to the cyclical utilization of the labour force."⁴ Some tests of my own similarly show a weak relationship between movements in real wages and the labor-capital ratio.⁵

A third source of evidence on the fixed proportions hypothesis (3) is obtained by examining trend-corrected capital-output ratios at successive cyclical peaks. Since the constraint $y_t \leq 1$ must be binding for capital to earn a positive return, the peak observed y_t should be (approximately) attained frequently. This prediction is contradicted by the levels of y_t , relative to trend, attained during and immediately following World War II, which far exceed any previous or subsequent values.⁶

⁴ The citation is from Section 7 of [2]. Bodkin uses the unemployment rate as an indicator of the cycle, and deflates money wages with a variety of price indexes. It should be mentioned that on postwar, quarterly U. S. manufacturing data Bodkin obtains results which (as he indicates) appear to be an exception to the results cited here.

⁵ These tests, based on data described in note 3, are reported in [7].

⁶ Output per unit of capital exceeded its secular (1890-1954) trend level by more than 20 percent in each of the years 1944-46, and in no other years of this period (using the Kendrick data in [5]).

To summarize, there appears to be no evidence of diminishing returns to labor in U.S. time series data, either from real wage movements or from direct estimation of production functions. On the other hand, the fixed factor proportions hypothesis cannot account for fixed output per unit of capital flows attained during wartime, or for the absence of countercyclical real wage movements.

II. An Alternative Model

The evidence summarized in the preceding section and its implications for the two standard production theories are familiar to empirical students of the business cycle. It is widely thought that the most hopeful route toward a resolution of these difficulties will be to incorporate variations in the rate of capital utilization into the theory. This can be done on two levels. First, some investigators have obtained "improved" empirical production functions (that is, have obtained labor elasticities closer to labor's share) by "correcting" measured capital stock for variations in utilization rates. Second, there have been attempts to reformulate the cost structure of firms to explain why less than full utilization of capital can be, at times, optimal. To account for the observations cited above, an adequate theory must clearly do both.

In this paper we shall pursue the view, advanced and studied in some detail by Marris, in which utilization is defined in terms of the fraction of hours per period (day, week) over which equip-

ment is operated.⁷ The rate of utilization, for fixed capital stock, will then depend on the rising schedule of wage rates as the firm moves from the most attractive to the least attractive hours of work. Certainly of the several available explanations for variations in rates of capital utilization, this view is the most relevant for secular movements. The basic reason that much of our capital is operated or occupied roughly 40 of 168 hours in a "normal" week is that people dislike night and weekend work, and the reason utilization in this sense has declined slowly throughout the century is that we have spent some of our increasing wealth on indulging this preference.⁸ The fact that both hours worked per week and shift work are pro-cyclical suggests that worker preferences may be important in understanding cyclical variation in capital utilization as well. To determine whether this is so, an explicit model linking preferences and utilization is developed in the remainder of this section.⁹

We think of time as divided into discrete periods, corresponding to the period of observation (e.g., quarters) where each period is of length one. Capital stock is fixed throughout the period at its beginning-of-period level, and output is sold (or added to inventory) at the end of the period. Time within the period is treated as the unit interval, with points ordered from most to least desirable for workers. From the point of view of the firm, these preferences are summarized in a schedule $w(s)$, $0 \leq s \leq 1$, of competitive wage rates for work at time s . This schedule will have an upward slope, reflecting observed premium pay for overtime, night, and weekend work.

The flow of output per unit of capital at the instant s , $y(s)$, depends on workers per unit of capital on duty at s , $x(s)$, according to the "instantaneous" production function, (1). The volume of output, per unit of capital, over the entire period of production, is then:

$$(4) \quad y = \int_0^1 y(s)ds = \int_0^1 f(x(s))ds.$$

Similarly, total man-hours per unit of capital are:

$$(5) \quad x = \int_0^1 x(s)ds,$$

⁷ See [8]. The view of utilization taken in this paper is essentially that of Marris, although we shall be concerned with a different set of implications. For alternative formulations, see Taubman and Wilkinson [11] and the "putty-clay" theories of Johansen [4] and Solow [10].

⁸ See the discussion by Lewis in [6].

⁹ The debt to Marris is acknowledged above. The following also uses the distinction between a rate and a volume of production in exactly the sense of Alchian [1].

and compensation of employees, per unit of capital, is:

$$(6) \quad wx = \int_0^1 w(s)x(s)ds.$$

An empirical production function, then, is a curve fitted to successive observations y_t and x_t (the left sides of the expressions (4) and (5)). Observed time series on money wages are observations on $(wx)_t$ divided by x_t (the expressions (6) and (5)). If the schedule $w(s)$ were flat, these variables would be related as discussed in the preceding section. The problem of this section is to relate these variables when $w(s)$ has the upward slope we observe.

Clearly, there will be no systematic relationship among these variables if the wage schedule $w(s)$ shifts "arbitrarily" from period to period. Accordingly, assume that $w(s)$ in period t is a base wage w_t times a premium schedule, $\phi(s)$ which remains stable over periods.¹⁰ Under this assumption, a firm which is scheduling labor optimally (or which is "on" its short-run cost curve) solves:

$$\max_{x(s) \geq 0} \int_0^1 f(x(s))ds$$

subject to:

$$(7) \quad \int_0^1 \phi(s)x(s)ds \leq c.$$

(That is, it maximizes output per unit of capital for a given total variable cost per unit of capital, cw_t .)

We first consider this problem under the fixed factor proportions hypothesis (3), or

$$(8) \quad f(x(s)) = \min[x(s), 1],$$

and:

$$(9) \quad \phi(s) \text{ is strictly increasing.}$$

From [13], Theorem 2'', a solution $x^o(s)$ exists. It is clear that $x^o(s)$ will take the value 1 on an interval $[0, u]$ and 0 on $(u, 1]$ for some number $0 \leq u \leq 1$. Since $\phi(s)$ is increasing, u is uniquely determined by:

$$(10) \quad \int_0^u \phi(s)ds \leq c, \quad \text{with equality if } u < 1.$$

The short-run product market implications of

¹⁰ There is some support for this assumption in the fact that some premium rates—notably overtime—are specified as a proportion of a base wage. About one-fourth of late shift workers receive a premium specified in percentage terms [12, p. 95]. A deeper analysis of this question would, of course, go behind the schedule $\phi(s)$ to the preference functions of labor suppliers.

this model are exactly as in the standard, neo-classical case. From (10) one obtains total variable cost as a function of u and w , $w_c(u)$, where $c(u)$ is strictly convex and satisfies:

$$(11) \quad c(0) = 0, \quad c(1) = \int_0^1 \phi(s) ds, \quad c'(u) = \phi(u).$$

From (4), output per unit of capital is $y = \int_0^1 \phi(s) ds = u$, so that $w_c \phi(y)$ is the marginal cost function.

From (5), total man-hours per unit of capital also equal u (and y) so that the observed output-labor input elasticity is predicted to be unity. With output per man-hour fixed over the cycle, the real wage and labor's share are the same variable. Their common value is given by the ratio of money compensation, $(w_t/u) \int_0^1 \phi(s) ds$, to output price (or marginal cost) $w_c \phi(u)$. The derivative of this ratio with respect to u (which is as good an index of the cycle as any other) is then:

$$\frac{1}{u} \left[1 - \frac{w}{p} \left(1 + \frac{u\phi'(u)}{\phi(u)} \right) \right].$$

Since w/p equals labor's share, it is between zero and one. On the other hand, $1 + u\phi'(u)/\phi(u)$ exceeds one. Hence the cyclical movement in the real wage rate is not restricted by the theory.¹¹ In summary, the theory just outlined is consistent with all of the evidence cited in the preceding section.

The short-run cost structure of the firm is essentially unchanged when the fixed proportions assumption is replaced by a neoclassical production function, as we next show. In place of (8), we require f to be twice differentiable and satisfy:

$$(12) \quad f(0) = 0, \quad f'(x) > 0, \quad f'(\infty) = 0, \quad f''(x) < 0.$$

Again, Theorem 2'' of [13] guarantees the existence of an optimal work schedule $x^0(s)$. The strict concavity of f and the convexity of the set of functions defined by $x(s) \geq 0$ and (7) assure that this optimum is unique. Since f is strictly increasing, (7) holds with equality at $x^0(s)$ and, further, there is a positive number λ such that, for all $0 \leq s \leq 1$,

$$(13) \quad f'(x^0(s)) \leq \lambda \phi(s),$$

with equality if $x^0(s) > 0$. To obtain the cost func-

tion of the firm, then, one must solve (7) and (13) for λ and $x^0(s)$.

Inverting (13) gives the optimal workforce at s as a function of $\lambda \phi(s)$: $x^0(s) = x(\lambda \phi(s))$. Then since the value of the objective function under $x^0(s)$ is output per unit of capital, y , we have:

$$y = \int_0^1 f[x(\lambda \phi(s))] ds.$$

This in turn may be solved for $\lambda = \lambda(y)$, where

$$(14) \quad \lambda(0) > 0, \quad \lambda(\infty) = 0, \quad \lambda'(y) < 0.$$

Then substitution into (7) yields total variable cost as a function of output:

$$(15) \quad w_c(y) = w_t \int_0^1 x[\lambda(y)\phi(s)] \phi(s) ds.$$

By differentiation and the application of (13), one verifies the familiar fact that marginal cost is:

$$w_c'(y) = \frac{w_t}{\lambda(y)}.$$

Thus, as in the fixed proportions case, the short-run product market implications of the model are identical to those of the standard neoclassical model discussed in the introduction.

This model implies a stable empirical "production function," obtained by substituting the optimal work schedule (which depends on output) into (5):

$$(16) \quad x = \int_0^1 x[\lambda(y)\phi(s)] ds.$$

The derivative of output with respect to labor input along this curve is readily calculated, but we can determine nothing about dy/dx except its sign. In particular, the elasticity $(x/y)(dy/dx)$ may be on either side of one, and bears no necessary relation to labor's share. Thus, this theory is consistent with the cyclical observations cited above, in the nearly vacuous sense that it is consistent with any stable empirical production function and with any cyclical pattern in real wages and labor's share.

III. Conclusions

The main point of this paper has been to examine two distinct margins along which a firm may increase its observed labor-capital ratio. The first, and most familiar, margin involves labor-capital substitution in the usual sense: at each point in time, one may work a fixed stock of equipment with more labor. The second corresponds to what is referred to as increasing the intensity of capital use: the fraction of the production period over which capital is

¹¹ This is true provided the elasticity $u\phi'(u)/\phi(u)$ is regarded as unobservable. The regression of $\log(w/p)$ on $\log(x)$ yields an estimate, according to the above theory, of the inverse of labor's share (about 1.5) less $1 + u\phi'(u)/\phi(u)$. Hence, a coefficient 0 on $\log(x)$ indicates an elasticity for $\phi(u)$ of .5. This number is not widely out of line with time-and-a-half for overtime or an 8 percent night shift differential [12, p. 86]. But to construct a careful test would involve the actual construction of the schedule $\phi(s)$, and the determination of the appropriate arc elasticity to compare against .5.

used is increased. Along both margins, increasing marginal cost is met: in the first case, due to diminishing returns in a productive sense; in the second, due to a rising schedule of premium wages as operations are extended to hours which workers regard as unattractive. These two sources of diminishing pecuniary returns, separately or in combination, have identical short-run product market implications, the same as those of the simple neoclassical theory of Section I. In contrast to the standard models, however, the models developed above are consistent with observed cyclical patterns in production and real wages. First, in common with the neoclassical theory, these models predict a stable relationship between output per unit of capital stock and man-hours per unit of capital stock. The elasticity of the former with respect to the latter need not, however, exhibit diminishing returns, even if the instantaneous production function is neoclassical. Thus both forms of the theory are consistent with the production function estimates reported in Section I. Second, neither of the models developed here requires a countercyclical pattern in average, real compensation per man-hour (the usual measure of "the" real wage rate). Thus the wage implications of the theory are consistent with the results of Bodkin [2] and others.

REFERENCES

1. ARMEN ALCHIAN, "Costs and Outputs," in Moses Abramovitz, ed., *The Allocation of Economic Resources: Essays in Honor of B. F. Haley* (Stanford Univ. Press, 1959).
2. RONALD G. BODKIN, "Real Wages and Cyclical Variations in Employment: A Reconsideration," *Canadian J. of Econ.*, Aug., 1969.
3. J. DURBIN, "Estimation of Parameters in Time Series Regression Models," *J. of the Royal Statist. Soc.*, Series B, 1960.
4. LEIF JOHANSEN, "Substitution Versus Fixed Production Coefficients in the Theory of Economic Growth: A Synthesis," *Econometrica*, Apr., 1959, pp. 157-76.
5. JOHN W. KENDRICK, *Productivity Trends in the United States* (Princeton Univ. Press, 1961).
6. H. G. LEWIS, "Hours of Work and Hours of Leisure," *I.R.R.A.*, 1956, pp. 196-207.
7. ROBERT E. LUCAS, JR., "Capacity, Overtime and Investment," unpublished Carnegie-Mellon research memorandum, Aug., 1969.
8. ROBIN MARRIS, *The Economics of Capital Utilization* (Cambridge Univ. Press, 1964).
9. EDMUND S. PHELPS, introductory chapter in E. S. Phelps et al., *The New Microeconomics in Employment and Inflation Theory* (W. W. Norton, 1969).
10. ROBERT M. SOLOW, "Substitution and Fixed Proportions in the Theory of Capital," *Rev. of Econ. Studies*, June, 1962, pp. 207-18.
11. PAUL TAUBMAN AND MAURICE WILKINSON, "User Cost, Capital Utilization and Investment Theory," *Int. Econ. Rev.* (forthcoming).
12. U.S. Dept. of Labor, Bur. of Labor Statis., *Wages and Related Benefits* (Part ii), Bul. No. 1530-87 (Washington, 1968).
13. MENAHEM E. YAARI, "On the Existence of an Optimal Plan in a Continuous Time Allocation Process," *Econometrica*, Oct., 1964, pp. 576-90.

DISCUSSION

ARNOLD C. HARBEGGER: It is in a sense embarrassing for me to comment on the paper by Professor Christensen, since its spirit has so much in common with some of my own earlier work. It is hard, I would think, for anybody, but especially difficult for me, to quarrel with an approach asking what are the effects of tax policy measures when consumers behave rationally and markets function as they should, in a model embodying the conditions of a full and general equilibrium.¹

In the analysis of tax policy, moreover, the most important effects to consider are, in my view, the long-term effects. They are what the economy has to live with indefinitely. "Announcement effects" are instantaneous and for that reason evanescent; the transition between announcement and the longer-term equilibrium is by definition transitory. The long run is precisely what Professor Christensen focuses on, and for the reasons given it is hard to fault him for that. His message, too, is a powerful one. In the final analysis the long-run effects of investment incentives on output will stem from their influence on the amounts of basic resources (labor and capital) which are voluntarily supplied to the market. Ultimately, it is how these tax stimuli affect the labor-leisure choice and the savings decision that will determine their effect upon output. Christensen notes that saving is likely to be insensitive to the rate of return (another assumption I have often made in the past) and produces two cases—one in which the counterpart of the reduction in revenue caused by the tax incentives is added borrowing from the household sector and the other in which the counterpart is reduced government investment—which result in opposite directions of movement of private investment.

Having agreed that Christensen's long run is the most fundamental level on which to analyze tax incentives, let me add that there is an intermediate sort of run about which some points of interest can be made concerning tax incentives to investment. Christensen's long run is so long and his model is so "real" that the money stock does not even appear in it, and monetary policy has no relevance for it. As a consequence he focuses only upon the fiscal counterparts of investment incentives. If one shifts the focus to allow for the possibility of less than full employment, the effects of fiscal incentives turn out to depend rather critically on the monetary policy being pursued, and the manner in which they so depend helps, I believe, to elucidate certain aspects of U.S. economic policy in the past decade or so.

Let us consider, at the one extreme, that in the absence of fiscal incentives to investment U.S. monetary policy would have been such as to keep the path through

time of real national income the same as it has in fact been in the presence of the tax incentives. In that case the incentives would have operated to raise interest rates above the level they otherwise would have reached and would have stimulated investment only to the degree that the higher interest rates stimulated saving. Believing the influence of interest rates on saving to be small, I judge that the major effect of the fiscal incentives, under the present assumption about monetary policy, was to shift investment out of areas not covered by the tax incentives and into areas to which those incentives applied. In point of fact, total investment and saving actually fell slightly as a percentage of GNP as one compares the periods 1955-61 (before the tax incentives) with 1962-67 (after the major policy shift to the investment tax credit and shortened Bulletin F lives), but residential construction fell from an average of 4.9 to 3.9 percent of GNP, while private domestic nonresidential investment rose from 10.4 to 11.1 percent of GNP.²

At the other extreme, one might consider that in the absence of the tax measures favoring investment, the monetary authorities would have caused the interest rate to behave through time in a fashion identical to what it in fact did in the presence of the incentives. In this case the incentives would by definition have no effect on the interest rate; the vehicle through which they worked would be national income. Income would rise, not only because of the impact-effect of increased spending (through moving along the demand curve for investment goods as the cost of capital was reduced), but also (at least so long as there was sufficient slack in the economy) because of the multiplier effects generated thereby.

In the intermediate case in which the quantity of money is assumed to follow the same path through time in the presence of the tax incentives as in their absence, the incentives affect both output and prices, since at any given time the incentives cause the *IS* curve to be higher, while (by assumption in the present case) the *LM* curve is stable.

In a sense, the message emerging from this exercise is the same as Christensen's: when tax incentives are analyzed in a general-equilibrium framework, the effects attributed to them will vary depending on the complementary policies pursued. In his case the complementary policies concerned how fiscal resources were obtained to replace the revenue loss which the incentive entailed. In my discussion I have implicitly ruled out added borrowing as the source of such replacement revenue, simply because it would so clearly operate to deter the private investment spending that the tax incentives were designed to stimulate. In my analysis above, I deal no further with the issue of the source of revenue replacement than is implied by the assumption that the *IS* curve shifts to the right (i.e.,

¹ These assumptions are basic in my "The Incidence of the Corporation Income Tax," *J. P. E.*, June, 1962; in my "Taxation, Resource Allocation and Welfare," in *The Role of Direct and Indirect Taxes in the Federal Revenue System* (Princeton Univ. Press, 1964), pp. 25-75; in my "The Measurement of Waste," *A. E. R.*, May, 1964, as well as other writings in the field of public finance.

² See my "Tax Stimuli and Investment Behavior," in Gary Fromm, ed., *Tax Incentives and Investment Spending* (Brookings Institution, forthcoming).

it is assumed that the revenue sources from which the lost funds are made up either do not operate to reduce the demand for investment, or, if they do so, do not have a negative effect upon desired investment which is stronger than the positive effect entailed in the tax stimuli). Making this very weak assumption about how the makers of fiscal policy choose among the alternatives upon which Christensen's analysis focuses, I have shown, in the context of *IS-LM* analysis, that the effects of tax incentives also depend critically on what monetary policies one assumes would have been pursued in their absence.

Some readers may be disturbed by the direction in which both Christensen's and my results lead—that one can say very little, *a priori*, about the effects of even so relatively simple a disturbance as the U.S. tax incentives to investment. But I am afraid that the multiplicity of possible outcomes to which we both allude is an almost inevitable aspect of policy analysis. Working in a general-equilibrium context even with very simple models, there are usually many relatively plausible ways of setting the *ceteris paribus* conditions under which the analysis is undertaken and of establishing what I have called above the counterparts of a given policy change. All of which means that it will be a long time, if ever, before economists will be able to reach a general professional accord, *ex ante*, about what the effects of any given changes are likely to be, even though they may agree on both the structure and the numerical parameters of the model to be employed.

VERNON L. SMITH: Economists are often intrigued by the apparent stability of the great ratios of economics either because they are interpreted as "laws of nature," or because they seem to belie the effectiveness of some cherished economic policy. The constancy of labor's share in the national product is the most famous counterpoint to the alleged efficacy of labor union power. Recently, in the face of anxiety over the rising threat of the military-industrial complex, it has been pointed out that since World War II defense spending has been a remarkably stable proportion of the national product. Now, *à la* Harberger and Christensen, we have the intransigence of domestic investment in proportion to the GNP.

One explanation of all this constancy—suggested half seriously many years ago by an eminent economist—might be found in the degree of cooperation and communication among the government staff members who compile aggregate statistics. Perhaps such statisticians use the ratios as a check on their calculations, and when the ratios depart from their expected constancy this sets forth a successful search for those errors which restore the desired levels.

But assuming that the aggregate data do reflect the true nature of the process there is a very simple explanation of the constancy of the investment-GNP ratio. It is, after all, consistent with the naive Keynesian investment multiplier for an economy whose consumption is proportional to income. The point is that the conclusion that tax policy has been ineffective in stimulating investment does not automatically follow from

the observation of a constant investment-GNP ratio. The observation may say more about the structure of the economy than the effectiveness of tax policy.

The important contribution of Christensen's paper is to provide a general equilibrium model that is sufficiently rich to permit at least a minimal evaluation of the allocative effects of tax policy. The major theoretical conclusion of Christensen's model is that the effect of investment tax incentives depends on the fiscal package.

I am not in basic disagreement with Christensen's approach and would like to turn this discussion to a somewhat different but related issue: the effect of tax policy on the direction rather than the aggregate of investment activity. Although there is probably no such thing as a neutral tax, I believe the laws governing depreciation for income tax purposes are particularly pernicious in distorting the direction of total investment activity. The concept of capital depreciation, illusive as it is to begin with, must in practice apply only to durable capital goods. But if we define a capital expenditure as any outlay today for the purpose of increasing future income, then it is clear that some of our most important forms of business investment are not in hardware but in research and development, the training and further education of technical, scientific, and managerial employees, "learning by doing," and advertising. All these categories of investment are fully expensed for corporate income tax purposes and are thereby given a subsidy of nearly 50 percent by comparison with investment in hardware. The economic consequences of this are that the tax depreciation laws are not likely to have the same effect on a railroad or steel manufacturer that they have on a pharmaceutical, cigarette, or management consulting firm whose investment in knowledge (including advertising) may be more important than their expenditures for hardware. When the 7 percent tax credit was introduced, some of this imbalance may have been corrected, but in a very haphazard and ambiguous way. It would seem much simpler and more sensible to permit all capital goods to be expensed for computing taxable income. The objective would not be to stimulate aggregate investment, even if this were a possible effect, but—for tax purposes—to treat all capital expenditures in the same manner.

There is considerable anecdotal evidence to suggest that the tax laws have important effects on decision making to the point of stimulating new institutions or industries devoted to what I have called "tax law arbitrage" activities. The leasing industry is perhaps the best known example. Before the 1962 liberalization the equipment lease contract was employed to achieve, in effect, a shortened write-off period for capital assets. This fact, by the way, could help to explain why the liberalization of depreciation did not seem to have any effect on the investment-GNP ratio. The 7 percent investment tax credit has encouraged leasing in a different way: Firm A has long-term capital requirements so large that it exceeds the total allowable credit under the law. Firm B has no significant long-term capital requirements, but the law imputes value to inventing requirements. So B has an incentive to supply the equipment needed by A under a lease contract. B owns the

equipment, gets the tax credit, and the resulting tax saving is shared with A through the rental price.

The social productivity of such contractual arrangements are rather suspect. My own conjecture is that the tax treatment of capital outlays may be more effective in fostering such artificial "arbitrage" activities than in stimulating aggregate investment in an economy that could seem to use a little less stimulation in any case. I think the 7 percent investment tax credit was ill-advised whether or not it is an effective stimulant, and I would suggest that we consider letting businesses treat capital goods as an ordinary expense, just as we now let them expense other outlays made in anticipation of increasing future earnings, and which increase net worth.

RALPH B. BRISTOL, JR.: As Professor Bischoff points out, nonresidential construction is not a common subject of study; yet it is an important and variable component of total output. If we look at the postwar movements in this investment sector, we see a strong boom in 1956-57 followed by seven years of essentially no growth. Then, starting in the first quarter of 1965, expenditures (in both current and constant dollars) increased more in four quarters than they had in the previous ten years. Subsequently, while expenditures have not fallen back to their levels of the early 1960's, they have not advanced any farther. I have been unable to find any investment function which, fit to a sample period ending in 1964 or earlier, can explain that boom. Also, if the sample period ends in 1965, every function I have tried fails to predict the abrupt cessation of that boom. I do not know whether any of Bischoff's equations could pass these tests or not, but I suggest them as an interesting challenge.

In his paper, Bischoff first builds a conventional distributed lag, stock adjustment model. As seems to be the universal practice, he assumes that desired stock is a function of expected or planned output. This last is, of course, approximated by a linear function of past levels of output. I would like to make two comments on this. First, would stock-adjustment models be as popular as they are if we described them the way we actually use them? It seems reasonable to say that desired future stock depends on some anticipated future flow, but to posit that the stock desired for next year is a function of last year's sales appears to assume pretty naive investors. Second, I would like to question the "output" variable: business gross product. Is this the best measure to use? I assume that at the micro level, sales are the desired variable, but when we aggregate across firms and industries, is business gross product the relevant figure?

To expand on his simple stock-adjustment model, Bischoff merely adds a rental cost of capital (*à la* Jorgenson) as an extra (linear) term. Deriving the rental cost of capital requires a discount rate. Rather than pick one *a priori*, Bischoff performs all calculations with ten alternatives. This disarming display of openmindedness (or ignorance) leads to the conclusion that the best discount rate involves the Standard & Poor dividend-price ratio for corporate stocks. This is in line with the

conclusion of other researchers. I have derived functions for producers' durable equipment and nonresidential construction based on Bischoff's earlier work on PDE, and the Federal Reserve-MIT model also used functions based on his work. In both cases, interest rates on corporate bonds were clearly the more relevant measure for PDE, while dividend rates were more relevant for structures. Why should fluctuations in the stock market have a significant effect on structures but not on equipment? Why are long-term bond rates highly relevant for equipment but not at all for structures? I would be glad if someone would provide me with a rationalization of this difference. I would like also to question the fact that the formulation is (mostly) in terms of "constant dollar" quantities. First, there is a "mechanical" effect in that Bischoff's *P/C* term contains the ratio of the private business deflator to the structures deflator. This has been on a downward trend throughout the sample period, meaning that the response of investment to changes in depreciation lives and taxes is less now than earlier. This may be true, but I wonder if it is not merely an artifact of the way we derive our price deflators. The other aspect of current versus constant dollar magnitudes is the tax rate, which, after all, applies to current dollar flows regardless of movements in deflators. Any divergence between current and constant dollar magnitudes indicates some degree of inflation and one of the likeliest results of inflation is the creation of capital gains in structures. I think the possibilities for converting regular income to capital gains and the tax treatment of capital gains should be included in any explanation of this area of investment. I am afraid I have nothing helpful to suggest concerning just how this should be done, but the existence of tax-sheltered income sources in construction is too well known to be ignored.

At the end of his paper, Bischoff concludes that he has developed a model which does a good job of explaining the fluctuations in nonresidential construction. Do we now understand what "causes" such investment expenditures? Can we use his model for forecasting future changes? Can we predict how such investment will respond to policy changes?

It seems clear that the investment booms of the mid-1950's and the mid-1960's both came about at the expense of residential construction. Visualize, if you will for a moment, somebody from the housing industry calculating regressions using the same independent variables, but with residential construction as the dependent variable. He would come out with the conclusion that if you want to restrain the economy (through the investment sector), you would carry out such measures as tax cuts, more liberal depreciation lives, and investment credits. Totally aside from multiplier effects, he would point out, these measures are strongly deflationary. (The implication would be clear: if you included multiplier effects, there would be disaster!) Well, there is no need to pursue the example any further, as the moral is clear: we are dealing with only partial models, and these can yield only partial explanations. I just cannot visualize what we are likely to end up with when we finally achieve more general equilibrium

models of the investment sector. If, as Bischoff's equations would indicate, the linkage between nonresidential construction and other investment is through stock market prices, I am afraid we are a long way from complete explanations.

I see that I have fallen into the usual discussant's trap: rather than discuss what the author has written, I have criticized him for a paper he did not write. My only defense is that Professor Bischoff has written a

fine paper which, by his willingness to experiment with alternative formulations, he has rendered impervious to general criticism. I felt his earlier work on PDE functions was excellent, and I have stolen his formulation for my own use. I intend to do the same with his present work on investment in structures.

If I have seemed critical, it is because I think that the time has come to move on beyond partial demand models.

MONEY WITHIN THE GENERAL FRAMEWORK OF THE ECONOMIC SYSTEM

INDICATORS OF MONETARY POLICY: THE ARGUMENTS AND THE EVIDENCE*

By MICHAEL J. HAMBURGER
Federal Reserve Bank of New York

During the past few years there has been considerable debate among economists over the most appropriate way of measuring the relative tightness or ease of monetary policy; that is, the problem of selecting variables to serve as indicators (or gauges) of the impact of monetary policy on economic activity.¹ Most observers are agreed that the variables chosen must be subject to Federal Reserve control, but this is as far as the consensus seems to go.

Through the years, measures of money market conditions—short-term interest rates, free reserves, and member bank borrowing from Reserve Banks—have been widely used as indicators, both inside and outside the Federal Reserve System. Such measures are regarded by many as important because they are believed to play a critical role in transmitting the effects of Federal Reserve actions to the economy.²

Recently, a growing number of economists have rejected money market condition variables as indicators of monetary policy. On the one hand, the monetarists [4] [7] [13] [14] [26] [27] take the view that movements in these variables do not

provide very useful information concerning subsequent movements in the variables of ultimate interest: employment, output, and prices. They contend that the total monetary impulse significantly shaping the behavior of the latter variables is best gauged by the growth rate of the money stock, and that the thrust applied by monetary policy is most usefully assessed in terms of the growth rate of the monetary base (high-powered money; i.e., total member bank reserves adjusted for changes in reserve requirements plus the public's currency holdings).³ The results obtained in a recent study by Andersen and Jordan [2] provide some support for this hypothesis. They find that for the period 1952 through 1968 quarterly first differences in GNP are highly correlated with current and lagged first differences in both the money stock and the monetary base.⁴

The monetarists are not alone, however, in rejecting money market condition variables as policy indicators. A number of other economists have reached the same conclusion, although for different reasons. They argue that the variables selected as indicators should be statistically exogenous; i.e., they should not be affected by current movements in income and other endogenous forces. The measure most often suggested by the proponents of this view is effective nonborrowed reserves—nonborrowed reserves adjusted for changes in reserve requirements (see, for example,

* For helpful criticism I am grateful to E. Gerald Corrigan, Richard G. Davis, Robert G. Link, Anton S. Nissen, Robert B. Platt, and other colleagues at the Federal Reserve Bank of New York. An earlier version of the paper was presented at the meeting of the Federal Reserve System Committee on Financial Analysis, Oct., 1969, Kansas City, and at seminars at the University of Rochester, the University of California (San Diego), and the Federal Reserve Bank of Philadelphia. Comments on those occasions were helpful. In addition, I wish to acknowledge the excellent research assistance of Susan K. Skinner.

¹ In this paper we do not consider the related and important problem of selecting policy targets. Targets of policy—as opposed to indicators—have been defined by Brunner and Meltzer [5] and Davis [10] as the variables the Federal Open Market Committee instructs the Manager of the System Open Market Account to hold within specified ranges during the intervals between its meetings. Alternative definitions are provided by Horwich and Hendershott [22] and Saving [28]. The need for targets and indicators is discussed in [6] [7] [10] [28] [29].

² The most recent statement of this view is presented by Smith [29]. In addition, see Starleaf and Stephenson [30].

³ Brunner and Meltzer define money as demand deposits plus currency, while Friedman and his associates extend the definition to include time and savings deposits at commercial banks.

⁴ Since quantitative indicators are generally expressed in terms of percentage growth rates it would have been preferable, for purposes of this analysis, if Andersen and Jordan had used logarithmic first differences rather than first differences of the variables in arithmetic form. Andersen and Jordan define money as demand deposits plus currency. In their earlier study Friedman and Meltzer [13] obtained results that were less favorable to the monetarist position. Using current and lagged quarterly first differences in the broader definition of the money supply for the period 1945–58 they were unable to explain more than 5 percent of the variance in the quarterly first differences in net national product.

[11] [15] [16] [21]). According to these economists all of the other commonly mentioned indicators—the money market condition variables, as well as the monetary base and the money stock—reflect the effects of both policy and nonpolicy influences and hence do not provide reliable (i.e., unbiased) measures of Federal Reserve actions.⁵

This paper examines some empirical evidence bearing on the various propositions discussed above. In Section I we test the hypothesis that effective nonborrowed reserves are less responsive to endogenous forces than other monetary variables. Section II takes up the other side of the question; that is, the extent to which movements in various measures of monetary influence provide reliable information on current and future movements in economic activity. The concluding section summarizes the results and discusses their implications.

I. *The Exogeneity Question*

The possibility of using econometric models to estimate the effects of macroeconomic policy has stimulated considerable interest in the variables used to represent government actions. If these variables are not statistically exogenous, biased estimates of the effects of policy are likely to be obtained. This holds, of course, regardless of whether one is estimating the parameters of a single reduced-form equation or those of a complete structural model. Critics of the Andersen and Jordan study [2] have therefore been particularly concerned over their treatment of the monetary base and the money supply as exogenous variables. The critics suggest that it would be more appropriate to use effective nonborrowed reserves, since the monetary authorities are, by and large, in a position to control the movements in this variable autonomously. But, as numerous writers have pointed out, there are many variables, including the money stock, that could be closely controlled by the monetary authorities (at least on a quarterly average basis), if they chose to do so.⁶

De Leeuw and Kalchbrenner argue that for any

measure of monetary influence to be exogenous it must be largely independent of current disturbances in the endogenous variables. This, they contend, requires: (1) that the Federal Reserve exert a "heavy and direct influence" over the variable (i.e., take it or some closely related magnitude as a policy target) and (2) that the variable does not respond contemporaneously to events in the private and other government sectors of the economy.⁷ Given that the Federal Reserve does not make policy independent of developments in the economy, it is quite difficult to devise variables that satisfy both of these conditions.

On the one hand, there appears to be fairly general agreement that during the period considered in most econometric studies, the Federal Reserve did not attempt to exert very much control over the growth rates of nonborrowed reserves, the money supply, or the monetary base, but instead focused its attention on money market conditions.⁸ Thus, monetary aggregates would not appear to satisfy the first of the De Leeuw and Kalchbrenner conditions. Guttentag argues, for example in discussing the period 1951–60, that although the behavior of the money supply and other "strategic" variables was always reported and often discussed at Federal Open Market Committee meetings, "the overwhelming impression gained from the Committee minutes is that these variables carry an effective weight close to zero. The basic reason is that typically the Committee does not hold firm views as to what constitutes appropriate behavior of these variables; within rather wide limits it is prepared to rationalize whatever behavior materializes. Thus, under the [Federal Reserve's] money market strategy, the value of strategic variables is, within wide limits indeterminate" [18, p. 14]. In addition, Guttentag reports that current data on nonborrowed reserves were not even reported to the Committee before May, 1960 [18, p. 27].

On the other hand, money market conditions variables do not satisfy the second condition, since there is ample evidence to suggest that the Federal Reserve adjusts the levels of its money market targets in response to changes in economic activity. The simple correlations between quarterly first differences in GNP and those in free reserves, borrowed reserves, and the three-month Treasury bill rate are -0.52 , $+0.51$, and $+0.51$, respectively (for the period 1953–68). Thus, none of the commonly mentioned indicators of mone-

⁵ This approach to the indicator problem is also discussed in Brunner and Meltzer [6], Davis [10], Kaufman [23], Saving [28], and Starleaf and Stephenson [30]. Brunner and Meltzer argued, at that time, that the monetary base (less member bank borrowing) is exogenous and is the "true" indicator of policy, if policy actions are restricted to open market operations [6, pp. 195–96]. Another exogenous policy variable proposed by Hendershott [20] and Horwich and Hendershott [22] is the neutralized money stock.

⁶ See Hendershott [20], Meek [25], Meltzer [27], and Smith [29].

⁷ De Leeuw and Kalchbrenner [11, pp. 6–8]. A similar set of conditions is proposed by Wood [31].

⁸ Among others see [1] [5] [12] [18] [24] [31]. For a conflicting view see [16] [17].

TABLE 1
PERCENT OF THE VARIANCE IN SIX MONETARY
INDICATORS EXPLAINED BY CURRENT MOVEMENTS
IN GNP: QUARTERLY LOGARITHMIC FIRST
DIFFERENCES, 1953-68

Indicator	1953-68	1953-60	1961-65	1966-68
<i>NBR</i>	.13†	.27†	—	.28*
<i>TR</i>	—	.01	.00	.01
<i>B</i>	.07*	—	.00	—
<i>M₁</i>	.21†	.18†	.05	—
<i>M₂</i>	.05*	—	.07	.01
<i>BC</i>	.02	—	.00	—

* Significant at the .05 level.

† Significant at the .01 level.

tary policy would appear to qualify as truly exogenous variables according to the De Leeuw-Kalchbrenner criterion.

If one wishes to use exogenous measures of monetary policy in econometric models, I can see little alternative but to proceed along the lines suggested by Wood. He argues that if policy-makers do not behave randomly with respect to economic events, then we must treat their instruments (e.g., nonborrowed reserves or the monetary base) as jointly dependent variables and introduce behavioral equations explaining the movements in those instruments [31, p. 135]. This line of reasoning suggests that it is the ultimate objectives of monetary policy (e.g., the desired rate of inflation) that should be treated as exogenous rather than the instruments used to achieve the objectives. Considering the importance generally attributed to exogenous policy variables by econometricians, there has been surprisingly little movement in this direction.

In the absence of such an analysis it is useful to take a look at some available evidence concerning the effects of endogenous forces on various monetary aggregates. Although such evidence does not in any sense provide a conclusive test of the exogeneity question, it does permit an examination of the often stated hypothesis that there is less statistical influence running from endogenous forces to effective nonborrowed reserves than to other indicators of monetary policy. (See [11] [15] [16] [21].) To begin with, we consider the relationship between current movements in income and six monetary aggregates. Table 1 shows the adjusted R^2 's for equations relating first differences in the logarithms of the monetary variables to contemporaneous first differences in the logarithm of GNP. The data are seasonally

adjusted quarterly averages for the period 1953-68 and three subperiods: 1953-60, 1961-65, and 1966-68. The latter were chosen to permit a test of Smith's conjecture that from 1961 through 1965 Federal Reserve policy was essentially passive in the sense that "causation ran mainly from income to the monetary base" [29, p. 12]. In addition to the base (*B*) the monetary variables included in the analysis are:

NBR: effective nonborrowed reserves (hereafter nonborrowed reserves)

TR: total reserves (effective nonborrowed reserves plus member bank borrowing)

M₁: the money supply (private holdings of currency and demand deposits)

M₂: the money supply plus time deposits

BC: bank credit⁹

Changes in logarithms are used to reduce the bias that might be introduced as a result of the strong common time trends in the first differences of the variables.¹⁰ Moreover, such changes correspond very closely to the units in which quantitative indicators are often expressed (percent changes). Dashes (—) are used in this and subsequent tables to denote those instances where the estimates of R^2 are negative; that is, where the percent of the variance explained by the equation is less than the loss in degrees of freedom.

A glance at the table should convince the reader of the general lack of statistical association between contemporaneous movements in income and the monetary variables—particularly total reserves, the monetary base, *M₂*, and bank credit. Nearly one-half of the 24 estimates of R^2 are less than or equal to zero, and 19 (80 percent) are less than 0.10. All five remaining estimates are associated with nonborrowed reserves and *M₁* and

⁹ The data for all monetary variables other than bank credit are averages of daily figures. For bank credit we use a three-month average of midmonth averages of bank credit at all commercial banks reported as of the last Wednesday of each month.

¹⁰ The possibility of such a bias is discussed in Andersen and Jordan [2] and Davis [9]. The nature of the problem may be illustrated by considering the behavior of the monetary base. The simple correlations among time ($t=1, \dots, 64$) and the arithmetic first differences in the base and GNP for the period 1953-68 are:

$$r_{\Delta B \Delta GNP} = 0.58 \quad r_{\Delta B} = 0.80 \quad r_{\Delta GNP} = 0.64$$

The strong trend components in ΔB and ΔGNP suggest that much of the observed correlation between these variables may merely be due to the fact that each of them has grown fairly steadily over time. The partial correlation between these variables holding t constant ($r_{\Delta B \Delta GNP \cdot t} = 0.15$) supports this view and indicates the importance of making some allowance for trend.

none occur during the period 1961–65. Thus, the evidence provides little or no support for the hypotheses that nonborrowed reserves are less affected by changes in income than other measures of monetary policy and that policy was essentially passive during the early 1960's.

It is interesting to note that the coefficient of GNP in the nonborrowed reserve equation (not shown in the table) is negative in each of the cases where the estimate of \bar{R}^2 is significant, suggesting that these variables tend to vary inversely.¹¹ While such a relationship may be desirable in terms of stabilization policy, it raises serious questions about the use of nonborrowed reserves as an exogenous variable in econometric models of the U.S. economy and about the estimates of the effects of monetary policy derived from such models (see, for example, the Brookings model and the FRB-MIT model).

The current value of GNP is only one of many endogenous variables that might be used to explain the movements in various indicators. Hence, it might be argued that much of the influence of economic activity on "money" is reflected by other variables, such as interest rates [9] [19].¹² To test this hypothesis changes in the logarithms of the six monetary aggregates (m) were regressed on current and lagged changes in the logarithms of GNP and the three-month bill rate (TBR). The general form of the equations is

$$(1) \quad \Delta \log m_t = a + b_0 \Delta \log GNP_t + b_1 \Delta \log GNP_{t-1} + c_0 \Delta \log TBR_t + c_1 \Delta \log TBR_{t-1}$$

Table 2 shows the estimates of \bar{R}^2 when only the current values of the independent variables are included in the equations; estimates for the complete equations are in Table 3.

The results are similar to those presented earlier in three respects. First, nonborrowed reserves are more closely correlated with the endog-

¹¹ For the full period 1953–68 the estimated regression is

$$\Delta \log NBR = .016 - .507 \Delta \log GNP \quad \bar{R}^2 = .131 \\ (3.24) \quad SE = .013$$

where the t -statistic is in parentheses. Additional evidence is provided by the cyclical behavior of nonborrowed reserves. From 1953 through 1968 the average annual rate of growth of effective nonborrowed reserves was 9.52 percent during NBER reference cycle downswings and 0.81 percent during upswings. In contrast, the growth rate of the monetary base was 2.06 in downswings and 2.28 in upswings, reflecting a very slight procyclical bias.

¹² Davis [9] also finds little if any feedback effect of GNP on money.

TABLE 2
PERCENT OF THE VARIANCE IN SIX MONETARY INDICATORS EXPLAINED BY CURRENT MOVEMENTS IN GNP AND THE TREASURY BILL RATE: QUARTERLY LOGARITHMIC FIRST DIFFERENCES, 1953–68

Indicator	1953–68	1953–60	1961–65	1966–68
<i>NBR</i>	.34†	.41†	.09	.23
<i>TR</i>	.11*	.12	—	—
<i>B</i>	.10*	—	.03	—
<i>M</i> ₁	.20†	.15*	.01	—
<i>M</i> ₂	.09*	—	.08	—
<i>BC</i>	.09*	.02	—	—

* Significant at the .05 level.

† Significant at the .01 level.

TABLE 3
PERCENT OF THE VARIANCE IN SIX MONETARY INDICATORS EXPLAINED BY CURRENT AND LAGGED MOVEMENTS IN GNP AND THE TREASURY BILL RATE: QUARTERLY LOGARITHMIC FIRST DIFFERENCES, 1953–68

Indicator	1953–68	1953–60	1961–65	1966–68
<i>NBR</i>	.39†	.48†	.09	.11
<i>TR</i>	.19†	.26*	—	—
<i>B</i>	.12*	—	.04	—
<i>M</i> ₁	.24†	.15	.01	.05
<i>M</i> ₂	.20†	.20	—	.24
<i>BC</i>	.17†	.12	—	—

* Significant at the .05 level.

† Significant at the .01 level.

enous variables than any other indicator.¹³ Second, during the period 1961–68 fluctuations in total reserves, the money supply, and bank credit appear to be largely independent of those in GNP and the bill rate. Finally, the monetary base behaves very much as might be expected if it were an exogenous variable: it is almost totally unrelated to current and past movements in GNP and the bill rate in each subperiod.

Thus, we find that while treatment of the money supply and the monetary base as exogenous variables may yield biased estimates of the effects of monetary policy on the economy, it seems unlikely, on the basis of the results presented here, that the use of nonborrowed reserves

¹³ A decrease in \bar{R}^2 from one table to the next indicates that the variable (or variables) added to the equation contribute less to the explanatory power of the model than they cost in terms of degrees of freedom.

would improve matters in any way, particularly during the 1960's. These findings are, of course, subject to some limitations. In particular, there is the possibility that all the relevant variables have not been included on the right-hand side of equation (1) and hence the R^2 's reported above may not provide unbiased estimates of the influence of endogenous forces on the monetary variables. However, it is not immediately clear what the missing variables might be and suggestions along this line would be most helpful.

II. *The Explanatory Power of Various Indicators*

We turn next to the question of which indicator provides the most reliable information on the future course of economic activity. One of the monetarists' principal assumptions is that the "monetary" effects of monetary policy dominate the "credit" effects. Hence, they argue that the total monetary impulse significantly affecting the economy is more appropriately measured by the growth rate of the money stock (defined either to include or exclude time deposits) than, say, by the growth rate of bank credit [4] [13] [27]. In addition, they suggest that the thrust applied by monetary policy should be assessed in terms of the growth rate of the monetary base rather than in terms of changes in its composition or the growth of individual components; e.g., nonborrowed reserves [4] [14] [26].

This view conflicts sharply with one that has been frequently expressed by Federal Reserve spokesmen and which is incorporated in many econometric models. According to the latter, the effect of borrowed reserves on bank credit, deposit expansion, and hence economic activity is different from the effect of nonborrowed reserves.¹⁴ Thus, a shift in the composition of member bank reserves away from nonborrowed reserves (e.g., a decrease in nonborrowed reserves holding total reserves constant) is often taken as an indication that monetary policy has become more restrictive.¹⁵

A test of the monetarists' hypotheses is provided by regressing changes in the logarithm of GNP on current and lagged changes in the loga-

rithms of the six quantitative indicators considered in the previous section (the results are presented in Table 6, below). The importance of the composition of member bank reserves may be evaluated by including current and lagged changes in the percent of reserves supplied through open market operations (the ratio of nonborrowed to total reserves $\Delta(NBR/TR)$) in the total reserve equation (see Table 4). In all but one case the structure of the lags is estimated by use of the Almon distributed lag technique, with the length of the lags chosen to maximize R^2 .¹⁶ In addition, an initial lag of one period is assumed where this contributes to the explanatory power of the equations. For the purposes of this analysis the sample period (1953-68) is divided into two subperiods: 1953-60 and 1961-68.

If the composition of member bank reserves is a useful measure of the thrust of monetary policy, changes in the percent of reserves supplied through open market operations should after some reasonable period of time begin to have a positive effect on the growth rate of GNP. The results presented in Table 4 provide no evidence at all of such an effect. Although most of the coefficients of the $\Delta(NBR/TR)$'s are not statistically significant, practically all of them are negative.¹⁷ On the other hand, the coefficients of the change in total reserves always have the anticipated signs and most of them are significant at the .05 level. Similar results are obtained when arithmetic first differences in GNP are regressed on current and lagged arithmetic first differences in nonborrowed reserves (ΔNBR) and borrowed reserves (ΔBR) (see Table 5). The borrowed reserves coefficients almost always exceed those for nonborrowed reserves, but the differences are rarely large and are never statistically significant. Moreover, for all practical purposes, the coefficients for the two subperiods are identical. Thus, the results suggest that it is changes in the total quantity of reserves available for deposit creation which influence the real sector of the economy; whether the additional reserves are supplied through open market operations or otherwise acquired appears to make no discernible difference.

As indicated above a comparison of the explanatory power of total reserves and the other

¹⁴ In answer to questions posed by the Commission on Money and Credit the Federal Reserve stated in 1960 that "shifts between policies of [monetary] ease and restraint often do not involve absolute changes in the total amount of reserves available to the banking system, but rather changes in the source of reserves as between reserve credit made available through open market operations, on the one hand, and loans and advances to member banks on the other" [8, p. 118].

¹⁵ See the interpretation of monetary developments in the latter half of 1968 in [3, pp. 3-28].

¹⁶ Due to programming limitations the Almon lag was estimated using a second degree polynomial rather than the fourth degree polynomial which has become popular.

¹⁷ In view of the problem of incorrect signs, it seemed appropriate to obtain unconstrained estimates of these coefficients. Hence, the lag structures associated with the $\Delta(NBR/TR)$'s were estimated without the use of the Almon procedure.

TABLE 4
REGRESSIONS EXPLAINING CHANGES IN THE LOGARITHM OF GNP:
QUARTERLY OBSERVATIONS, 1953-68
(*t*-values in parentheses)

Period	Constant	$\Delta \log TR$				$\Delta (NBR/TR)$			
		<i>t</i> -1	<i>t</i> -2	<i>t</i> -3	<i>t</i> -4	<i>t</i>	<i>t</i> -1	<i>t</i> -2	<i>t</i> -3
1953-68	.803 (4.56)	.162 (1.32)	.271 (4.26)	.281 (3.59)	.190 (2.94)	-.416 (2.73)	-.060 (0.38)	.018 (0.10)	-.103 (0.73)
		$R^2 = .500$		$SE = .761$		$DW = 1.66$			
1953-60	.660 (2.52)	.318 (1.43)	.407 (2.89)	.384 (2.49)	.248 (2.07)	-.380 (1.66)	-.049 (0.21)	.005 (0.02)	-.133 (0.66)
		$R^2 = .448$		$SE = .971$		$DW = 1.70$			
1961-68	1.01 (3.07)	.100 (0.63)	.211 (2.20)	.231 (2.03)	.161 (1.77)	-.240 (0.98)	-.004 (0.16)	.000 (0.00)	-.151 (0.59)
		$R^2 = .253$		$SE = .398$		$DW = 1.70$			

five quantitative indicators is presented in Table 6. The equations estimated are of the form

$$\Delta \log_t GNP = a + \sum_{i=1}^{n_2} b_i \Delta \log m_{t-i}$$

where n_1 denotes the initial lag and takes the values zero or one, and n_2 varies from 4 to 20.¹⁸

¹⁸ A table containing the estimates of n_1 , n_2 and the sum of the b_i 's is available upon request from the author.

Generally speaking, the results are impressive. The estimates of R^2 for total reserves, the monetary base, M_1 , and bank credit are considerably higher than those in Tables 1 through 3. This suggests that monetary policy as measured by the growth rates in these variables has an important and predictable effect on economic activity. This is particularly true for total reserves and bank credit since only lagged values of these variables are included in the equations. The evidence

TABLE 5
REGRESSIONS EXPLAINING ARITHMETIC FIRST DIFFERENCES IN GNP:
QUARTERLY OBSERVATIONS, 1953-68
(*t*-values in parentheses)

Period	Constant	ΔNBR					ΔBR				
		<i>t</i>	<i>t</i> -1	<i>t</i> -2	<i>t</i> -3	<i>t</i> -4	<i>t</i>	<i>t</i> -1	<i>t</i> -2	<i>t</i> -3	<i>t</i> -4
1953-68	2.8 (3.6)	1.7 (0.7)	7.6 (5.7)	10.4 (6.8)	10.1 (6.0)	6.6 (5.4)	6.0 (1.3)	10.6 (5.3)	12.2 (7.4)	11.0 (5.6)	7.0 (4.6)
		$R^2 = .671$		$SE = 3.72$			$DW = 1.61$				
1953-60	2.8 (2.3)	0.8 (0.2)	6.7 (2.0)	9.5 (2.6)	9.4 (2.5)	6.2 (2.5)	6.5 (0.9)	9.8 (2.4)	10.8 (2.5)	9.6 (2.1)	5.9 (1.9)
		$R^2 = .425$		$SE = 4.26$			$DW = 1.63$				
1961-68	4.2 (2.7)	0.7 (0.2)	6.7 (3.6)	9.6 (4.2)	9.5 (3.8)	6.3 (3.5)	1.4 (0.2)	8.5 (2.4)	12.0 (4.9)	11.7 (4.1)	7.7 (3.5)
		$R^2 = .575$		$SE = 3.26$			$DW = 1.78$				

TABLE 6

PERCENT OF THE VARIANCE IN GNP EXPLAINED BY
CURRENT AND LAGGED MOVEMENTS IN VARIOUS
INDICATORS: QUARTERLY LOGARITHMIC FIRST
DIFFERENCES, 1953-68

Period	<i>NBR</i>	<i>TR</i>	<i>B</i>	<i>M</i> ₁	<i>M</i> ₂	<i>BC</i>
1953-68	.32	.44	.34	.39	.28	.39
1953-60	.23	.43	.30	.26	.14	.31
1961-68	.22	.31	.21	.31	.43	.45

does not, however, provide very much support for the specific indicators proposed by the monetarists. Bank credit explains a larger percentage of the variance in GNP than either definition of the money supply during both subperiods. In addition, the total reserves variable always performs better than the monetary base and in all but one case is more closely related to GNP than is *M*₁. The results for the second subperiod (1961-68) are, however, heavily dependent upon the use of the Almon distributed lag technique. If unconstrained regressions are used the estimate of *R*² for the narrowly defined money supply equation rises to 0.49 while all the other estimates fall. Thus, although a strong case can be made for measuring the relative tightness or ease of monetary policy by the growth rates of various quantitative indicators, the empirical evidence is not very helpful in suggesting which quantity other than nonborrowed reserves is the most important one to look at.

III. Conclusion

Despite their obvious limitations the results presented above are useful in pointing up some of the gaps in our understanding of monetary economics. First, they do not support an assumption that has been incorporated in most of the recent econometric models of the U.S. economy; namely, that effective nonborrowed reserves are more exogenous than most other monetary variables. Second, they raise serious questions about the appropriateness of differentiating between changes in borrowed and nonborrowed reserves in estimating the effects of Federal Reserve actions. Finally, they suggest that income is more closely related to items on the asset side of bank balance sheets—credit—than to items on the liability side—money. Although I personally find it difficult to accept a bank credit theory of the monetary mechanism, the evidence would appear to put the monetarists under some obligation to provide an explicit analytical framework expressing

their conception of the mechanism, something they have thus far failed to do. The latter finding should also be of interest to the architects of large-scale econometric models, since they too have paid very little attention to the supply of financial assets to banks; i.e., the public's demand for liabilities.

REFERENCES

1. L. C. ANDERSEN, "Money Market Conditions as a Guide for Monetary Management," in A. D. Entine, ed., *Monetary Economics: Readings* (Belmont, Calif., 1968), pp. 226-45.
2. L. C. ANDERSEN AND J. L. JORDAN, "Monetary and Fiscal Actions: A Test of their Relative Importance in Economic Stabilization," *Review*, Fed. Res. Bank of St. Louis, Nov., 1968, pp. 11-21.
3. Board of Governors of the Federal Reserve System, *Annual Report, 1968* (Washington, D.C., 1969).
4. K. BRUNNER, "The Monetary Fiscal Dilemma," *Bul. of Bus. Res.* (Ohio State Univ.), June, 1969.
5. K. BRUNNER AND A. H. MELTZER, *The Federal Reserve's Attachment to the Free Reserves Concept* (Washington, D.C., 1964).
6. ———, "The Meaning of Monetary Indicators," in G. Horwich, ed., *Monetary Process and Policy: A Symposium* (Homewood, Ill., 1967), pp. 187-217.
7. ———, "Targets and Indicators of Monetary Policy," in K. Brunner, ed., *Targets and Indicators of Monetary Policy* (San Francisco, 1970).
8. Commission on Money and Credit, *Board of Governors of the Federal Reserve and the United States Treasury Department: Answers to Questions from the Commission on Money and Credit* (Englewood Cliffs, N.J., 1963).
9. R. G. DAVIS, "How Much Does Money Matter? A Look at Some Recent Evidence," *Monthly Review*, Fed. Res. Bank of New York, June, 1969, pp. 119-31.
10. ———, "Measuring Monetary Policy." Paper presented at the Eleventh Annual Forecasting Conference, American Statistical Association, New York, Apr., 1969.
11. F. DE LEEUW AND J. KALCHBRENNER, "Monetary and Fiscal Actions: A Test of Their Relative Importance in Economic Stabilization—Comment," *Review*, Fed. Res. Bank of St. Louis, Apr., 1969, pp. 6-11.
12. W. G. DEWALD, "Free Reserves, Total Reserves and Monetary Control," *J.P.E.*, Apr., 1963, pp. 141-53.
13. M. FRIEDMAN AND D. MEISELMAN, "The Relative Stability of Monetary Velocity and the Investment Multiplier in the United States, 1897-1958," in Comm. on Money and Credit, *Stabilization Policies* (Englewood Cliffs, N.J., 1963), pp. 165-268.
14. M. FRIEDMAN AND A. SCHWARTZ, *A Monetary History of the United States, 1867-1960* (Princeton, N.J., 1963).

15. L. E. GRAMLEY, "Guidelines for Monetary Policy—the Case Against Simple Rules." Paper presented at the Financial Conference, National Industrial Conference Board, New York, Feb., 1969.
16. E. M. GRAMLICH, "The Role of Money in Economic Activity: Complicated or Simple?" *Business Economics*, Sept., 1969, pp. 21–26.
17. ———, "The Usefulness of Monetary and Fiscal Policy as Discretionary Stabilization Tools." Paper presented at the Conference of University Professors, American Bankers Association, Milwaukee, Sept., 1969.
18. J. M. GUTTENTAG, "The Strategy of Open Market Operations," *Q.J.E.*, Feb., 1966, pp. 1–30.
19. P. H. HENDERSHOTT, "Neutralization of the Money Stock," *Review*, Fed. Res. Bank of St. Louis, forthcoming.
20. ———, *The Neutralized Money Stock* (Homewood, Ill., 1968).
21. ———, "A Quality Theory of Money," *Nebraska J. of Econ. and Bus.*, Autumn, 1969.
22. G. HORWICH AND P. H. HENDERSHOTT, "The Appropriate Indicators of Monetary Policy," in D. P. Jacobs and R. T. Pratt, eds., *Savings and Residential Financing, 1969 Conference Proceedings* (Chicago, 1969), pp. 33–52.
23. G. G. KAUFMAN, "Indicators of Monetary Policy: Theory and Evidence," *Nat. Bank. Rev.*, June, 1967, pp. 481–91.
24. M. W. KERAN AND C. T. BABB, "An Explanation of Federal Reserve Actions (1938–68)," *Review*, Fed. Res. Bank of St. Louis, July, 1969, pp. 7–20.
25. P. MEEK, "Can the Desk Hit Aggregate Monetary Targets." Speech given at the Forum on Finance, Fed. Res. Bank of New York, June, 1969.
26. A. H. MELTZER, "The Appropriate Indicators of Monetary Policy," in D. P. Jacobs and R. T. Pratt, eds., *Savings and Residential Financing, 1969 Conference Proceedings* (Chicago, 1969), pp. 11–31.
27. ———, "Controlling Money," *Review*, Fed. Res. Bank of St. Louis, May, 1969, pp. 16–24.
28. T. R. SAVING, "Monetary-Policy Targets and Indicators," *J.P.E.*, Suppl., Aug., 1967, pp. 446–56.
29. W. L. SMITH, "A Neo-Keynesian View of Monetary Policy." Paper presented at "The Monetary Conference," Fed. Res. Bank of Boston, June, 1969.
30. D. R. STARLEAF AND J. A. STEPHENSON, "A Suggested Solution to the Monetary-Policy Indicator Problem: The Monetary Full Employment Interest Rate," *J. of Fin.*, Sept., 1969.
31. J. H. WOOD, "A Model of Federal Reserve Behavior," in G. Horwich, ed., *Monetary Process and Policy: A Symposium* (Homewood, Ill., 1967), pp. 135–66.

THE CHOICE OF OPTIMAL INTERMEDIATE ECONOMIC TARGETS*

By ROBERT HOLBROOK *and* HAROLD SHAPIRO
University of Michigan

I. Introduction

In the most general (nonoperational) sense, the goal of macroeconomic policy is clear and undisputed. It is to keep the economy as close as possible to some desired path of evolution. Further, there is probably general agreement that the determination of the economy's optimal path is "keyed" to the ultimate targets of full employment, stable growth, and equitable income distribution, subject to certain constraints regarding the balance of payments and price stability. Unfortunately, conflicts among these various objectives often exist (their precise nature depending upon the structure of the economy) which serve to limit the set of feasible paths for the evolution of the economic system. When this set of feasible paths has been defined, the selection of the optimal one becomes a matter of social choice. Thus, in a world where we have complete knowledge of the economic structure and full information on the current economic situation, where social preferences are revealed and where we have accurate forecasts of future exogenous elements affecting the economy, the determination of the optimal economic policy is a relatively straightforward matter, and the issues of intermediate targets and indicators do not arise.

Unfortunately, the determination of economic policy usually proceeds in an environment characterized by a number of important uncertainties. First, there are uncertainties about the structure of the economy. Second, knowledge about the current economic situation is only partial, as information concerning the current state of many critical variables (including the ultimate target variables) is available only after a considerable lapse of time. Finally, forecasts of the magnitude of future exogenous influences on the economic system have wide confidence intervals. In this environment policy-makers require some current guide to aid them in the continual adjustment of economic policy. The concept of intermediate targets has been developed to fill this need.

In principle, the actual response of policy-makers to this situation has often been quite rational, even if not optimal. They have generally proceeded by choosing some intermediate target which is readily available without information lags, responds quickly to variations in the policy instruments, and is thought to be related in a known way to some or all of the ultimate targets. The intermediate target then serves, in the absence of information on the current state of the ultimate targets, as one of the principal guides to the continual adjustment of current policy. If the intermediate target chosen is an endogenous variable of the system (e.g., interest rates) as opposed to a policy instrument (e.g., reserve base), the need for a distinct policy indicator, a variable which will properly isolate (indicate) the influence of current economic policy alone on the intermediate target variable, also arises, but is not the subject of the paper.

The present paper considers the problem of selecting an appropriate intermediate target to guide the adjustment of monetary policy in the presence of uncertainty. A relatively simple model is formulated and an optimal strategy is derived for the selection of an appropriate intermediate target, given certain assumptions about the desired levels of the ultimate targets. Our model does not allow for the full range of uncertainties that plague the policy-maker in the real world, but the stochastic elements introduced do yield some new insights into the general nature of the problem. Further, the analysis specifies the set of empirical information necessary to formulate actual policies. Our model differs from previous attempts to specify optimal economic strategies (see [1] [5]) in its emphasis on the selection of an optimal intermediate economic target. This emphasis is necessitated in our view by the policy-maker's uncertainty about the current values of at least some of the ultimate targets. In a non-stochastic model, of course, it makes no difference which intermediate target is selected. Once stochastic elements are introduced, however, the ultimate target variables themselves become random variables whose distribution depends on the particular settings of the intermediate economic targets. The remainder of the paper analyzes the selection of optimal intermediate targets for monetary policy in this environment.

* This study was assisted by a grant from the Canada Council. We are indebted to our colleague Professor R. Porter, for helpful comments on an earlier draft. The authors are also indebted to Robert Greig, Richard Hokenson, Steve Marston, and Harvey Rosen for their assistance. Responsibility for errors remains with the authors.

Our main conclusion is that optimal policies depend not only on the structure of the economy but on the particular nature of the uncertainties facing the policy-maker. Section II develops the analysis in terms of a simple *IS-LM* curve model, while Section III applies this analysis to an aggregative econometric model of the U.S. economy, to provide an example of the selection of an optimal intermediate target variable.

II. The Model

As a preliminary vehicle for our analysis we use a highly aggregative demand-determined *IS-LM* curve model with constant prices. Although so simple a model overlooks many important mechanisms in an actual economy, it is adequate to illustrate our analysis and under certain conditions can be considered a fair approximation of reality. The model consists of consumption and investment functions, demand and supply functions for money, and the identities necessary for closure. All equations are expressed in real terms. The model can be written as follows:

- (1) $C = A_1 + bY$
- (2) $I = A_2 + qY$
- (3) $Y = C + I + X$
- (4) $M^d = H + kY + sr$
- (5) $M^s = U + vr + wB$
- (6) $M^d = M^s = M$

where

C = Consumption¹

I = Gross Investment

X = Government Expenditure + Net Exports

Y = GNP

r = Interest Rate

M^d, M^s = Real Balances demanded and supplied, respectively

B = Real Monetary Base

The autonomous variables in the above model (A_1, A_2, H , and U) are portmanteau variables which are functions of all lagged and other exogenous variables. For the policy-maker, these variables are subject to two major types of uncertainty (randomness): First, he only has estimates of the coefficients connecting these variables to the rest of the economic system. Second, the values of the unlagged exogenous variables themselves may be subject to forecast error. For example, one might argue that the discount rate should help determine the autonomous term U

in the money supply equation (5). This would require a forecast of its value during the period in question, a forecast that is likely to be made with error. We also assume that all other coefficients in the above model (1)–(6) are represented only by estimates of their true value. Thus, we view each of the coefficients and autonomous elements in the system as being a random variable with a mean equal to its estimated (or forecast) value.

We assume that the policy-maker's ultimate target is a certain desired level of output (Y^*)², but that he operates with a one-period lag on information concerning this target variable (and, of course, concerning C and I as well). In addition, he has available estimates of b, q, k, \dots , etc., and forecasts of A_1, A_2, X, H , and U . His problem, then, is to select that intermediate target which will enable him "best" to achieve his ultimate objective. He can choose among three such intermediate targets in our model, the interest rate (r), the money stock (M), and the monetary base (B), any one of which he is assumed to be able to control without error.

The first issue is to choose the criterion by which the optimal intermediate target will be selected. For purposes of illustration we rank alternative policies by their ability to minimize the expected squared deviation of the actual value of output (Y) from the desired value of output (Y^*)³. Next, the above model (equations (1)–(6)) is solved for the reduced form equation for income (Y). Other reduced form equations of the system are unnecessary for the analysis, as only income and desired income appear in the loss function. The choice of the intermediate target determines the nature of the reduced form equation. For example, with an interest rate target, the interest rate is an exogenous variable and will appear in the reduced form, but with a money supply target, the interest rate is endogenous and must be solved out of the equations.⁴ The reduced form equations for income are given alternatively as (7), (8) or (9), depending on whether an interest rate, money

¹ This implies that all ultimate targets can be expressed in terms of a desired level of output.

² The policy-maker is assumed to select the value of each intermediate target variable such that the expected value of income is equal to desired income, and then to choose among the intermediate target variables that one which minimizes the expected squared deviation of actual from desired income. This is not equivalent to choosing that value of each intermediate target variable which minimizes the squared deviation of actual from desired income. The quadratic loss function used here has a number of interesting properties, and is fully discussed by Theil [5] and Hirschleifer [2].

⁴ We overlook the possibility that policy actions themselves are endogenous to the system.

³ For a more precise definition of these variables see Table 1 below.

supply or reserve base target is chosen.

$$(7) \quad Y = \frac{1}{1-b} [(A_1 + A_2 + X) + q\bar{r}]$$

$$(8) \quad Y = \frac{1}{(1-b) + \frac{kq}{s}} \left[(A_1 + A_2 + X) - \frac{q}{s} (H - M) \right]$$

$$(9) \quad Y = \frac{1}{(1-b) + \frac{qk}{s-v}} \left[(A_1 + A_2 + X) - \frac{q}{s-v} (H - U - wB) \right]$$

We can now evaluate our expected loss function $[E(L) = E(Y^* - Y)^2]$, for each potential intermediate target variable. We present the complete derivation only for the case of an interest rate target. The remaining cases (money supply and monetary base targets) are derived in an analogous fashion, and we simply present the corresponding solutions for them.

Using the reduced form solution for income, (7), the policy-maker derives that value for the rate of interest (his intermediate target variable) which he estimates will generate the desired level of income, as follows:

$$(10) \quad Y^* = \frac{A + q\bar{r}}{1-b},$$

where $\hat{A} = (\hat{A}_1 + \hat{A}_2 + X)$ and where the symbols $\hat{}$ and $\bar{}$ refer to estimated values and policy tools, respectively. The true value for income, however, is:

$$(11) \quad Y = \frac{A + q\bar{r}}{1-b},$$

and the difference between them can be written as:

$$(12) \quad Y^* - Y = \left(\frac{A}{1-b} \right) - \left(\frac{A}{1-b} \right) + \left[\left(\frac{q}{1-b} \right) - \left(\frac{q}{1-b} \right) \right] \bar{r}$$

Since some terms on the R.H.S. of (12) contain ratios of random variables, these terms are expanded in Taylor series around their means (discarding second and higher order terms) so that their expectations can be evaluated. This procedure yields the following expression:

$$(13) \quad Y^* - Y = \frac{(A - A)}{1-b} + \frac{(q - q)\bar{r}}{1-b} + \frac{(q - b)Y}{1-b}$$

In order to evaluate the loss function we square (13) and take its expectation, obtaining:

$$(14) \quad \begin{aligned} E(Y^* - Y)_1^2 &= E(L_1) \\ &= \left[\frac{dY}{dA} \right]_F^2 [\sigma_A^2 + Y^2 \sigma_b^2 + \bar{r}^2 \sigma_q^2 \\ &\quad + 2(Y\sigma_A \hat{b} + \bar{r}\sigma_A \hat{q} + Y\bar{r}\sigma_b \hat{q})] \end{aligned}$$

where

σ_j^2 = variance of j
 σ_{ij}^2 = covariance of i and j
 $dY/dA|_F = 1/1-b$ (the multiplier for autonomous expenditure when the interest rate is held constant)

For convenience, this can be expressed in matrix form as:

$$(15) \quad \begin{aligned} E(L_1) &= \left[\frac{dY}{dA} \right]_F [1, \bar{r}, Y] \\ &\quad \begin{bmatrix} \sigma_A^2 & \sigma_A \hat{b} & \sigma_A \hat{q} \\ \sigma_A \hat{b} & \sigma_b^2 & \sigma_b \hat{q} \\ \sigma_A \hat{q} & \sigma_b \hat{q} & \sigma_q^2 \end{bmatrix} \begin{bmatrix} 1 \\ \bar{r} \\ Y \end{bmatrix} \left[\frac{dY}{dA} \right]_F \\ &= Q_1 Z_1 \Sigma_1 Z_1' Q_1' \end{aligned}$$

where the subscript 1 refers to an interest rate target.

Loss functions for a money supply or a monetary base target can be derived with the same technique, and each yields a solution of the form:

$$(16) \quad E(L_i) = Q_i Z_i \Sigma_i Z_i' Q_i'$$

where

$$i = \begin{cases} 1 & \text{for an interest rate target} \\ 2 & \text{for a money supply target} \\ 3 & \text{for a monetary base target} \end{cases}$$

and

$$\begin{aligned} Q_2 &= \left(\frac{dY}{dA} \right)_M, - \frac{dY}{dM} \\ Q_3 &= \left(\frac{dY}{dA} \right)_B, - \frac{1}{\bar{w}} \frac{dY}{dB} \\ Z_2 &= \begin{pmatrix} 1 & r & Y & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & r & Y \end{pmatrix} \\ Z_3 &= \begin{pmatrix} 1 & r & Y & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & r & Y & -1 & -r & -B \end{pmatrix} \end{aligned}$$

Σ_1 = Variance-covariance matrix of the estimated coefficients in the order: A, q, b, H, s, k .

Σ_2 = Variance-covariance matrix of the estimated coefficients in the order: $A, q, b, H, s, k, U, v, w$.

In principle, all that remains to be done is to evaluate these functions and determine that intermediate target which minimizes the expected loss. This process is carried out in Section III (where the model is modified slightly to incorporate lags), but before proceeding with this task we

will use (15) and (16) as they stand to obtain certain simple qualitative results concerning the relative efficiency of interest rate and money supply targets.⁵

If we simplify the problem by assuming that the covariance terms are zero, the expected loss for an interest rate target is:

$$(17) \quad E(L_i) = \left(\frac{dY}{dA} \right)^2 [\sigma_A^2 + r^2 \sigma_q^2 + Y^2 \sigma_f^2]$$

where the term in square brackets can be thought of as a measure of the variability of the commodity demand sector (*IS* curve). The expected loss for a money supply target is

$$(18) \quad E(L_m) = \left(\frac{dY}{dM} \right)^2 [\sigma_M^2 + r^2 \sigma_q^2 + Y^2 \sigma_f^2] + \left(\frac{dY}{dM} \right)^2 [\sigma_H^2 + r^2 \sigma_s^2 + Y^2 \sigma_k^2]$$

where the second term in square brackets is essentially a measure of the variability of the monetary sector (*LM* curve).

We can compare the expected losses under these alternative policy regimes with the following expression and rule:⁶

$$(19) \quad V = \frac{\left[\left(\frac{dY}{dA} \right)^2 - \left(\frac{dY}{dM} \right)^2 \right]}{\left(\frac{dY}{dM} \right)^2} \cdot \left(\frac{\sigma_A^2 + r^2 \sigma_q^2 + Y^2 \sigma_f^2}{(\sigma_H^2 + r^2 \sigma_s^2 + Y^2 \sigma_k^2)} \right)$$

Rule: $\begin{cases} \text{If } V > 1 \text{ use money supply target.} \\ \text{If } V < 1 \text{ use interest rate target.} \end{cases}$

In the limiting cases of: (1) an interest inelastic demand function for money ($s=0$), (2) an interest inelastic commodity demand schedule ($q=0$), or (3) a liquidity trap, the results are clear. In the first case a money supply target is preferred while in the latter two cases any type of monetary policy is impotent. In all other cases, V will be positive, and its value will depend on the size of the multipliers and on the relative variability of the commodity and monetary sectors. *Ceteris paribus*, the more unstable is the commod-

ity demand sector relative to the monetary sector the more likely is a money supply target to be preferred. For example, if the only instability were in the commodity demand schedule, a money supply target would be preferred to an interest rate target. Here, the use of a money supply target permits stabilizing movements in interest rates whenever income deviates from its desired value; no such stabilizing effect operates if an interest rate target is chosen instead. If the only instability were in the money demand schedule, the interest rate target would be preferred. In this case the use of a money supply target would set up destabilizing movements in interest rates causing income to deviate from its desired value, while an interest rate target would permit the attainment of the desired level of income without error. If we are faced with both types of instability, optimal policy depends critically on the relative magnitudes of the uncertainties involved, as illustrated in (19).

It might appear that a simple way out of the dilemma posed by the selection of intermediate targets would be to have the policy-maker follow one such target most of the time, but to watch the other target(s) as well, and adjust policy so as to maintain it (them) within some predetermined interval around its (their) predicted value(s). For example, the rule might be to maintain a certain interest rate as long as the money supply grows at some rate between 2 and 6 percent, but if these targets turn out not to be compatible, then the money supply growth target becomes binding. This would be an entirely satisfactory strategy if the only source of uncertainty were the commodity demand schedule, since any deviation in the monetary growth rate from its predicted value would indicate an unexpected shift in that schedule. This shift would result in a deviation in income from its predicted value if the interest rate target is maintained, but a switch to a money supply target would permit the automatic stabilization process to come into play. (Of course, the same result would have been obtained if a money supply target had been followed from the beginning.) Suppose, however, that the only instability were in the money demand equation. Then deviations from the predicted rate of monetary growth would not affect income as long as the interest rate target were maintained, but a switch to a monetary growth target would cause income to deviate from its desired value. (This was just the problem that the choice of the interest rate target was supposed to overcome.) Uncertainty surrounds our predictions of both schedules, of course, and there is no reason to suppose that a policy of watching several targets and reacting to

⁵ Precise results require estimates of all the parameters involved, forecasts of the exogenous elements, and estimates of the elements of the variance-covariance matrices in each of the relevant expressions (15) or (16).

⁶ This rule is exactly equivalent to determining whether the expression $[E(L_i) - E(L_m)]$ is greater or less than zero.

where $X = Y - (C + I)$. This function is sufficiently simple to permit the derivation of the required loss function, while at the same time having a reasonable degree of explanatory power.

The monetary sector is represented by the following equations,

$$(21) \quad r_t = D + fY_t + gM_t \quad (\text{money demand})$$

$$(22) \quad M_t = U + vr_t + wB_t \quad (\text{money supply})$$

where the demand equation has been normalized with respect to the interest rate to permit its use in a three stage least squares procedure.⁷

when selecting the optimum value of his intermediate monetary target.⁸ It is this lack of knowledge about current income which opens up the problem of intermediate targets in this model; if he knows the value of Y_t , he will always optimize by selecting an interest rate target.

Using the same solution technique as before, the expected value of the loss function can be derived for each of the three policy options. These have the form:

$$(23) \quad E(L_t) = Q_t Z_t \Sigma_t Z_t' Q_t' + \sigma_x^2$$

where the matrices are constructed as follows:

$$Q_1 = \left[1, \frac{dY_{t+1}}{dY_t} \right]$$

$$Q_2 = \left[1, \frac{dY_{t+1}}{dY_t} \Big|_M, \frac{1}{g} \frac{dY_{t+1}}{dM_t} \right]$$

$$Q_3 = \left[1, \frac{dY_{t+1}}{dY_t} \Big|_B, \frac{1}{gw} \frac{dY_{t+1}}{dB_t}, \frac{1}{w} \frac{dY_{t+1}}{dB_t} \right]$$

$$Z_1 = \begin{bmatrix} 1 & X_{t-1} & C_{t-1} & I_{t-1} & Y_t & r_t & 0 \\ 1 & X_{t-2} & C_{t-2} & I_{t-2} & Y_{t-1} & r_{t-1} & 1 \end{bmatrix}$$

$$Z_2 = \begin{bmatrix} 1 & X_{t-1} & C_{t-1} & I_{t-1} & Y_t & r_t & 0 & 0 & 0 & 0 \\ 1 & X_{t-2} & C_{t-2} & I_{t-2} & Y_{t-1} & r_{t-1} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & Y_t & M_t \end{bmatrix}$$

$$Z_3 = \begin{bmatrix} 1 & X_{t-1} & C_{t-1} & I_{t-1} & Y_t & r_t & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & X_{t-2} & C_{t-2} & I_{t-2} & Y_{t-1} & r_{t-1} & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & Y_t & M_t & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & r_t & B_t \end{bmatrix}$$

Σ_1 = Variance-covariance matrix of coefficients in the order: a, e, m, n, b, q, X

Σ_2 = Variance-covariance matrix of coefficients in the order: $a, e, m, n, b, q, X, D, f, g$

Σ_3 = Variance-covariance matrix of coefficients in the order: $a, e, m, n, b, q, X, D, f, g, U, v, w$

The policy-maker is viewed as trying to choose that value of r_t , M_t , or B_t which will result in the smallest deviation of Y_{t+1} from its desired value, (Y_{t+1}^*). We estimate the target level of income (Y_{t+1}^*) by substituting the actual values of r , M , or B observed during the sample period into the appropriate reduced form expressions for Y_{t+1} . In the economy represented by (20) to (22), once period $t+1$ has arrived, the policy-maker no longer can influence Y_{t+1} , but only Y_{t+2} . During period t he is in possession of past values of all variables, but he does not know the true value of Y_t (current income), which becomes available only at the end of period t . He can, however, forecast Y_t with (20), and he will use this forecast value

⁷ In principle, the terms a , D , and U should be the sum of the constant term and the residuals for each equation. The procedure used in this empirical example ignores the residuals, and uses only the estimates of the constant terms and their standard errors. The effect of this omission is unclear.

The recursivity of the model permits the estimation of the real and monetary sectors separately, and makes all covariance terms between them equal to zero. In addition, we assume a zero covariance between the estimate of X and the other coefficients.

The values of the coefficients, and of the relevant variances and covariances, are shown in Table 1. These values, together with the actual values of X , C , I , r , M , and B over the sample period, were used to generate estimates of the expected loss functions for each quarter. These loss functions were estimated under two alternative assumptions regarding the policy-maker's ability to forecast public expenditures (X_t). First we evaluated the various loss functions assuming X_t was forecast through the use of a simple linear time trend (experiment A). Second we assumed

⁸ Our procedure for forecasting public expenditures (X_t) is described below.

perfect foresight regarding estimates of X_t (experiment B). Experiment A, by building in a rather crude forecast of X_t , insures a certain amount of unexpected shifting (instability) of the commodity demand schedule (IS curve). Experiment B is designed to eliminate this source of uncertainty.

The actual evaluation and ranking of the alternative loss functions for each of the three possible targets and in each of the two experiments produced a remarkably consistent result. In each quarter of the sample and postsample period (1967-I-1968-IV), the money supply target succeeded in minimizing the loss function.⁹ A monetary base target was only slightly inferior, while an interest rate target always performed most poorly. This result is not surprising for experiment A, which uses so crude a forecasting model for X_t , and thus builds a considerable amount of instability into the IS curve. We did not anticipate, however, that the ranking of the alternative intermediate targets would be unchanged in experiment B, despite the elimination of a major source of forecast error. We suggest two (not mutually exclusive) explanations for this result. First, it may be due to the misspecification of the commodity demand sector, resulting in larger than necessary errors in the forecast of current and future values of the target variable. Second, it may reflect an inherent instability in the commodity demand sector, not matched by a similar degree of instability in the monetary sector. Whatever the underlying cause, our results lead us to a tentative conclusion that the monetary policy-maker can best achieve a target level of income if he uses the money supply as his intermediate target.¹⁰

⁹ The estimated size of the loss function is much smaller for experiment B than experiment A, of course, but the policy ranking is the same in both cases.

¹⁰ It might be argued that our procedures bias the results in favor of the money supply target, since the assumption that the policy-maker can control this intermediate target variable without error is untenable. The money supply is influenced by a variety of forces which would make the attainment of a precise value

IV. Conclusion

Our conclusion in favor of a money supply target is based on two simplifying assumptions; namely, that the policy-maker is concerned only with achieving a target level of income, and that our model, (20)-(22), is the appropriate characterization of the U. S. economy. Neither of these assumptions is really satisfactory, however, and much more sophisticated analysis must be undertaken before final conclusion can be reached. Our simple loss function entirely ignores the problems of prices, growth, distribution, balance of payments, etc., while the policy-maker must keep all of these in mind. Our model, too, omits the supply side entirely, and thus lacks any price and distribution mechanism. Far better and more complete forecasting models than ours are available to the policy-maker, and he should take advantage of them. Thus, the analysis of intermediate targets must move toward the evaluation of more complete loss functions and more elaborate forecasting models. Until this is done, our results can, at best, only be taken as suggestive.

Our conclusion, however, is consistent with the casual observation that no one has yet blamed the failure of the tax surcharge on an unexpected shift in the demand function for money.

REFERENCES

1. W. BRAINARD, "Uncertainty and the Effectiveness of Policy," *A.E.R.*, May, 1967, pp. 411-25.
2. J. HIRSHLEIFER, "Investment Decisions Under Uncertainty-Choice-Theoretic Approaches," *Q.J.E.*, Nov., 1965, pp. 509-36.
3. J. KAREKEN, "The Optimum Monetary Instrument Variable" (Univ. of Minnesota, mimeographed).
4. W. POOLE, "Optimal Choice of Monetary Policy Instrument in a Simple Stochastic Model," *Q.J.E.*, forthcoming.
5. H. THEIL, *Optimal Decision Rules for Government and Industry* (Rand McNally, 1964).

impossible. Even if we were to accept this view, however, and decide that a money supply target is inferior to an interest rate target for this reason, it is still true that the monetary base scores substantially higher than the interest rate in the above experiment.

IMPLICATIONS OF FOUR ECONOMETRIC MODELS FOR THE INDICATORS ISSUE*

By RICHARD ZECHER
University of Chicago

Monetary economists concerned with the indicators of monetary policy issue are like dermatologists in that their patient seems incapable of either dying or getting well. The indicators issue cannot be expected to die so long as the two sets of prime indicator candidates—interest rates and monetary stocks—frequently yield conflicting information about the stance of monetary policy. If it is unlikely that the patient will die a quiet death, some consider it even more unlikely, in the near future of course, that he will be cured to everyone's satisfaction through the concentrated application of evidence. Rather than presenting additional hypotheses and evidence, this paper is concerned with revealing the implications of some existing hypotheses for the indicators issue. These hypotheses include four large-scale econometric models referred to here as FRS-MIT [5], FRS-CHI-MIT [13], Brookings [8], and Ando and Goldfeld [1]. Being hypotheses of unknown quality, these models are not likely to either kill or cure the patient, but they do contribute to the debate in several interesting ways.

Detailed discussions of the indicators issue are available in the literature [2] [3] [16], so only a brief summary is offered here: (1) An indicator of monetary policy is a variable which yields information about the thrust of current monetary policy actions on a goal variable in the present or some future time period. Commonly mentioned indicator candidates are prices and quantities observed in money markets. (2) In comparing the quality of two indicator candidates, one is said to be superior if it is more likely to yield true and less likely to yield false information than its competitor. Stated differently, an indicator of monetary policy should be relatively insulated from shocks other than those generated by monetary policy actions. These conditions are set out rigorously in the next section. (3) An ideal indicator, which reveals the separate, quantitative effect of policy actions on the goal variable in the present and in all future time periods, is differentiated from a less than ideal indicator, which may reveal only the direction of policy effects with some probability of being correct. According to the models mentioned above, the open market policy instrument (bank reserves or monetary base) approaches being an ideal indicator if no other policy

instruments are concurrently used. This assertion is also elaborated in the following section.

It is expedient and I believe correct to divide the participants in the indicators debate into price or interest rate watchers and quantity watchers. Although price watchers tend to be those stressing Keynesian income-expenditure theories and quantity watchers tend to be those stressing quantity theories of national income determination, there is no logical relationship between the choice of a theory and the choice of an indicator. For instance, even a perfectly constant ratio of money to income by itself implies nothing for the indicators issue (the money stock may be independent of monetary policy actions). Similarly, a theory with an explicit interest rate transmission mechanism from monetary to expenditure sectors does not imply that interest rates are the best indicator of the separate effect of monetary policy actions on expenditure variables. Noting that these are logically separate issues aids in explaining some of the implications of the models.

In deriving the implications of the models for the indicators issue, the field of indicator candidates is reduced to one price, the Treasury Bill rate r , and one quantity, total privately owned bank deposits M .¹ Two further simplifications are (1) of all the exogenous shocks which could make r or M yield false indicator information, only past changes in the open market instrument and current changes in fiscal policy represented by government expenditures are considered² and (2) price watchers are assumed to interpret increases in r as an indication of tight policy and decreases in r as indicating loose policy, while quantity

* The author gratefully acknowledges his indebtedness to Professors Karl Brunner and Allen Meltzer, and to members of the Money and Banking Workshop at the University of Chicago for their comments on earlier drafts of this paper.

¹ Two reasons for choosing r and M are the frequency with which they are mentioned as indicator candidates and the ease in obtaining the necessary data for these relative to other indicator candidates.

² These models have 40, 60, or even more exogenous variables, particularly if lagged variables are all treated as exogenous or predetermined. Choosing only three lagged changes in the open market variable and only the current change in government expenditures reflects my interpretation of factors most frequently discussed in the indicators debate.

TABLE 1

CHANGE IN NOMINAL INCOME DUE TO \$1.0 BILLION
INCREASE IN MONETARY POLICY INSTRUMENT*

Quarter	FRS- MIT	FRS- CHICAGO- MIT	Brookings	Ando- Goldfeld
1	0.7	1.2	-0.8	1.9
2	2.1	3.7	1.3	2.2
3	4.3	7.2	1.6	2.3
4	6.7	10.8	2.9	2.7

* The Policy Instrument is unborrowed reserves (Brookings and FRS-MIT), total reserves (FRS-Chicago-MIT), and total reserves plus currency Ando-Goldfeld.

SOURCES: Coefficients for FRS-MIT and FRS-Chicago-MIT are from solutions reported by Kaufman and Laurent in [13]. Brookings coefficients are from solutions reported by Fromm and Taubman in [8]. Ando and Goldfeld coefficients are from solutions reported in [1].

watchers are assumed to view decreases in M as tight policy and increases as loose policy.³ The four models are required to answer the question, are lagged monetary policy actions and current fiscal policy actions more likely to cause r or M to yield false information about the stance of current monetary policy?

I. Empirical Issues

Empirical issues separating price watchers from quantity watchers can be concisely derived from three reduced form equations of a general equilibrium model. Equations (1), (2), and (3) express the nominal interest rate r_t , nominal money stock M_t , and nominal income Y_t at time t as functionals of current and past monetary policy actions summarized by the nominal base B_t , current and past fiscal policy actions summarized by nominal government expenditures F_t , and all other predetermined forces current and past O_t , where t refers to current and past time periods back to period $t-n$.

$$(1) \quad r_t = f(B_t, B_{t-1}, \dots, B_{t-n}, F_t, \dots, F_{t-n}, O_t, \dots, O_{t-n}) + e_1$$

$$(2) \quad M_t = g(B_t, \dots, B_{t-n}, F_t, \dots, F_{t-n}, O_t, \dots, O_{t-n}) + e_2$$

³ In addition to this manner of assessing the information content of r and M , a second difference criterion is mentioned later in the text. Some observers may argue that it is the level of r that indicates tight or loose policy, and if they care to define ranges for "high" and "low" rates, they can give a third interpretation to the results in Tables 1-5.

$$(3) \quad Y_t = h(B_t, \dots, B_{t-n}, F_t, \dots, F_{t-n}, O_t, \dots, O_{t-n}) + e_3$$

The strong price watcher position on indicators, where Y is the policy goal, is summarized by two conditions:

$$(4) \quad (\partial r_t / \partial B_t) / (\partial Y^* / \partial B_t) \doteq k,$$

where k is a negative constant, and

$$(5) \quad |(\partial r_t / \partial B_t) dB_t| \gg |(\partial r_t / \partial X) dX|,$$

where X may equal any argument in functional f other than B_t .

Y^* is the sum of changes in Y due to a unit change in B after four quarters have elapsed. Stated otherwise, $\partial Y^* / \partial B_t$ is the sum of partial derivatives of the k functional with respect to the first four arguments in that functional, B_t , B_{t-1} , B_{t-2} , and B_{t-3} . Thus condition (4) asserts that a change in B will be associated with a current change in r , and a four quarter change in Y , that stand in a roughly constant proportion. A weaker statement of condition (4) asserts that k is a variable which, however, may only take negative values. Condition (5) is one way of asserting that any observed change in r is very likely due to changes in the stance of monetary policy. Taken together, (4) and (5) imply that the effect of current policy actions on income is expansive if r falls.

The quantity watchers' indicators position is represented by rewriting (4) and (5), inserting M everywhere r appears.

$$(6) \quad (\partial M_t / \partial B_t) / (\partial Y^* / \partial B_t) \doteq k',$$

where k' is a positive constant (strong assertion) or a variable taking only positive values (weak assertion), and

$$(7) \quad |(\partial M_t / \partial B_t) dB_t| \gg |(\partial M_t / \partial X) dX|,$$

TABLE 2

CHANGE IN NOMINAL INCOME DUE TO AN INCREASE
IN GOVERNMENT EXPENDITURES, BILLIONS
OF CURRENT DOLLARS*

Quarter	FRS- MIT	FRS- CHICAGO- MIT	Brookings†	Ando- Goldfeld
1	8.4	7.6	3.7	5.0
2	11.8	9.9	4.9	8.5
3	13.5	10.9	6.0	10.3
4	15.1	11.9	5.0	11.5

* \$3.8 billion for Brookings; \$5.0 billion for all others.

† Durable goods only.

SOURCES: See Table 1.

where X may equal any argument in functional g other than B .

Quantity watchers challenge the price watchers' position in part by attacking the validity of conditions (4) and (5). High and rising interest rates, they contend, are frequently associated over extended periods of time with rising and accelerating income in business cycle upturns, and particularly in such periods as 1965-66 and 1968 and part of 1969. Low and falling interest rates are associated with cycle downturns, a striking example being 1930-33. These observations, of course, do not refute the price watchers' position since monetary policy effects during these periods could have been as indicated by interest rates while other, opposing factors dominated changes in income. Thus quantity watchers concentrate their efforts on refuting condition (5).

Condition (5) is false, they contend, even if only the effects of monetary policy actions on interest rates are considered [6]. This argument identifies three effects of monetary policy actions on interest rates. First is the wealth, portfolio, liquidity, or simply current relative price effect. This is the widely recognized effect of an expansive policy move that increases base money and the money stock and reduces interest rates. The second or income effect of the policy move is distributed over several succeeding quarters. Increases in income due to the original expansive policy move tend to increase interest rates. The third effect follows at some interval behind the income effect and may be called a Fisherian price anticipations or intertemporal relative price effect. This effect tends to increase interest rates by the amount of anticipated inflation. In sum, quantity watchers argue that a rise in the current, observed nominal rate of interest is as likely to reflect past expansionary policy as current restrictive policy.

Price watchers have likewise concentrated their attacks on the second condition (7) of the quantity watchers' position [4] [10] [11] [12] [15]. Changes in relative yields on money and non-money assets, they argue, make the response of monetary stocks to current monetary policy actions unpredictable. In addition to this current relative price effect, they note an income effect on current changes in monetary stocks which may be due to past monetary policy actions, current or past fiscal actions, or other factors. These two forces generally tend to change monetary stocks in opposite directions. In cyclic upturns, for instance, nonmoney yields tend to rise relative to yields on money; thus reducing observed monetary stocks, while rapidly rising income increases money demand. Price watchers contend that the

net effect of these two forces is likely to be large relative to the separate effect of current monetary policy actions on monetary stocks, and hence that changes in observed monetary stocks are not dominated by current policy actions.

Arguments relating to the effects of anticipated price-level changes on nominal money stocks are not in evidence in the literature. The effects of anticipated price-level increases on nominal money stocks are not clear, except to the extent monetary authorities impose effective yield ceilings on bank deposits reducing nominal money stocks.

II. *Implications of the Models*

Conditions (4) and (6) are met by all four models because for these particular solutions of the models, increases in B decrease r and increase M and Y . Several additional solutions of all these models except Ando and Goldfeld have been made by the model builders and others. These solutions, which use different initial conditions (time periods), unambiguously indicate that the models are consistent with at least the weak versions of both conditions (4) and (6), and hence do not on this basis discriminate between the price and quantity watchers' positions.

These solutions also imply that the open market instrument, whether base money or a reserve figure, would be a superior if not quite ideal indicator. That is, $\partial Y / \partial B$ is stable in all models in the sense that it never changes sign and changes value only slowly and predictably over time.

Since all four models imply both conditions (4) and (6), it is possible to concentrate exclusively on the models' implications for conditions (5) and (7). Recall that condition (5) asserts the price watchers' position that changes in r are dominated by current changes in B , and condition (7) asserts the quantity watchers' position that changes in M are dominated by current changes in B . This must be taken to mean that each group feels that it is very unlikely that past monetary policy actions, current fiscal policy actions, or any other exogenous shocks have large effects (large relative to the effects of current monetary policy effects) on their indicator candidate.

The models' implications for the two indicator positions are derived by first defining what is meant by true and false indicator information, and then showing under what conditions r and M yield true and false indicator information. The r is said to yield false indicator information if $\text{sign } \Delta r = \text{sign } \Delta B$, and otherwise to yield true information. M yields false indicator information if $\text{sign } \Delta M \neq \text{sign } \Delta B$, and otherwise M yields true information. Past changes in B or current changes in

TABLE 3
CHANGES IN THE SHORT-TERM INTEREST RATE DUE TO MONETARY
AND FISCAL POLICY IN BASIS POINTS*

Quarter	Monetary†				Fiscal‡			
	FRS-MIT	FRS-CHI-MIT	Brook-ings	Ando-Goldfeld	FRS-MIT	FRS-CHI-MIT	Brook-ings	Ando-Goldfeld
1	-114.7	-210.9	-148.1	-83.6	26.5	29.6	3.9	8.7
2	-27.3	-24.5	-57.6	-54.1	19.4	22.4	3.7	17.6
3	-43.9	-55.9	-47.9	-43.0	22.1	26.1	3.5	24.0
4	-39.0	-35.4	-41.5	-11.9	23.4	27.8	1.8	27.9

* Definitions in footnotes of Tables 1 and 2.

† Billion dollar change in Open Market Instrument.

‡ \$3.8 billion change for Brookings; \$5.0 billion change for all others.

SOURCE: See Table 1.

fiscal policy F can make Δr , ΔM , or both yield false indicator information. We now turn to the conditions under which these events would occur according to each of the four models.

The initial conditions of this experiment are that effects of all predetermined variables other than current changes in B on M and r cancel to zero in the current and all subsequent periods. Let the monetary authorities increase B by \$1.0 billion, having effects on r and M reported in Tables 3 and 4. In the period following the initial change in B , changes in B can be divided into ranges making M a true and r a false indicator, ranges making r a true and M a false indicator, and ranges making both r and M either true or false indicators. These outcomes for the FRS-

MIT model are summarized on the following line. For $\Delta B > \$0.77$ billion,

	r true M true	r true M false	r false M true	r true M true	ΔB billion dollars
FRS-MIT		-0.28	0	0.77	

both r and M correctly indicate expansive monetary policy, and for $\Delta B < \$-0.28$ billion both r and M correctly indicate tight monetary policy (for effects of B on Y and definitions of variables for each model, see Table 1 and footnotes to Table 1). If ΔB falls between $\$-0.28$ and $\$0.00$ billion, r rises correctly indicating tight policy while M increases incorrectly indicating loose monetary policy. If $\$0.00$ billion $< \Delta B < \$0.77$

TABLE 4
CHANGES IN A MONEY STOCK DUE TO MONETARY AND FISCAL POLICY,
BILLIONS OF CURRENT DOLLARS*

Quarter	Monetary†				Fiscal‡			
	FRS-MIT	FRS-CHI-MIT	Brook-ings	Ando-Goldfeld	FRS-MIT	FRS-CHI-MIT	Brook-ings	Ando-Goldfeld
1	4.7	9.1	3.2	1.8	1.0	0.7	0.1	0.0
2	6.0	10.0	3.6	2.8	1.3	0.7	0.1	-0.2
3	7.0	10.5	4.1	3.7	1.7	0.8	0.2	-0.4
4	8.0	11.1	4.6	4.0	2.1	0.8	0.2	-0.7

* Money stock equals privately owned time and demand deposits; definitions of policy actions in footnotes of Table 1.

† See footnotes * and † to Table 3.

SOURCES: See Table 1.

billion, r is a false and M a true indicator.

Similar ranges for the other models are derived from the information in Tables 3 and 4. As before $\Delta B = \$1.0$ billion the first quarter, and ranges of ΔB in the second quarter which make r and M true and false indicators are derived. If no policy action is taken in the quarter following

FRS-	r true	r true	r false	r true	
CHICAGO-	M true	M false	M true	M true	ΔB
MIT					billion
		-0.10	0.0	0.88	dollars
Brookings	r true	r true	r false	r true	ΔB
	M true	M false	M true	M true	billion
		-0.13	0.0	0.61	dollars
Ando-	r true	r true	r false	r true	ΔB
Goldfeld	M true	M false	M true	M true	billion
		-0.56	0.0	0.36	dollars

the initial $\Delta B = \$1.0$ billion, similar ranges can be derived for the third and fourth quarters. Since the lagged responses of both r and M decay or approach new stable values very quickly, the ranges which discriminate between r and M also become very small in later quarters. In general, the discriminating ranges approach zero quickly in later quarters, and the r false range more quickly than the M false range.

There are two clear and important patterns in these discriminating ranges. All of the only r true ranges necessitate a reversal in the direction of change in B , while all only M true ranges do not require such a reversal, and all but one of the cases favoring r are for the third and fourth quarters where all discriminating ranges are small. Thus to the extent that B tends to increase every quarter, the probability of a false signal from M due to effects of lagged ΔB approaches zero. To the extent that B tends to decelerate from time to time, the probability of a false signal from r is increased. The second pattern suggests that to the extent that B tends to increase every quarter, the small discriminating ranges in the third and fourth quarters are of diminishing importance relative to the new shocks. In sum, these models imply that past policy actions are more likely to cause r than M to yield incorrect information about the current thrust of policy on income.

To further reveal the models' implications for the effect of past monetary policy actions on current values of r and M , an arithmetic example based on three possible monetary policies is presented. The three policies used to generate the numbers in Table 5 are (1) $\Delta B = \$1.00$ billion per

quarter, (2) $\Delta B = \$1.00$ billion in the first quarter, and ΔB is sufficient in later quarters to make $\Delta r = 0.0$, and (3) $\Delta B = \$1.00$ billion in the first quarter, and is sufficient in later quarters to make $\Delta M = \$2.00$ billion per quarter. Policies (1) and (3) are monetary rules not unlike those suggested by Friedman [7], while policy (2) could result from a desire to maintain present money market conditions, or a desire to conform to the assignment rule [14], or to some constant growth rule.

Policy (1) is designed to yield a constant impact of monetary policy on expenditures. By the direction of change criteria, both M and r correctly indicate the direction of policy actions for every period and for all four models. Using the acceleration or second difference criteria, interest rates fail to indicate the correct direction of policy in two out of three periods for models one and two, and for all three periods for models three and four. The money stock correctly indicates loose or easy policy in all cases. The ranges over which changes in r and M vary reveal that for these models past monetary policy actions are relatively more important in determining current changes in interest rates than in determining current changes in the money stock.

Policy (2) is to hold r at the level reached by the policy action $\Delta B = 1.0$ in period t . The sole purpose of policy actions in periods $t+1$, $t+2$, and $t+3$ is to offset the lagged policy effects on r . These latter changes in B have a large expansive effect on income which, of course, is not revealed by changes in r . Changes in money, by the direction criteria, correctly indicate expansive policy in every case and, furthermore, are large when changes in B are large and small when changes in B are small. Using the acceleration criteria, changes in money incorrectly indicate tight policy in all but three cases.

The last policy rule holds changes in M to $\$2.0$ billion per quarter after an initial shock. Here the interest rate incorrectly indicates the thrust of policy half the time using the direction of change criteria, while changes in money provide correct information in every case. There is also a very close association between the sizes of changes in M and B , and hence in F . Using the acceleration criteria, the interest rate gives false signals one-half and the money stock one-third of the times.

According to these three arithmetic examples, the models imply that when only the effects of current and past monetary policy actions are considered the interest rate is more likely to yield false information about current policy impulses than is the money stock. Other examples can be constructed which show the money stock giving

TABLE 5
THREE MONETARY POLICIES AND THEIR EFFECTS ON
THE BILL RATE r AND MONEY STOCK M^*

Period	Δr (basis points)				ΔB (billion \$)				ΔM (billion \$)			
	Model†				Model†				Model†			
	1	2	3	4	1	2	3	4	1	2	3	4
Rule: $\Delta B = 1.0$ per quarter												
t	-114	-211	-148	-84	1.0	1.0	1.0	1.0	4.7	9.1	3.2	1.8
$t+1$	-27	-25	-58	-54	1.0	1.0	1.0	1.0	6.0	10.0	3.6	2.8
$t+2$	-38	-46	-48	-43	1.0	1.0	1.0	1.0	7.0	10.5	4.1	3.7
$t+3$	-34	-26	-42	-12	1.0	1.0	1.0	1.0	8.0	11.1	4.6	4.0
Rule: $\Delta r = 0$ after first quarter												
t	-114	-211	-148	-84	1.0	1.0	1.0	1.0	4.7	9.1	3.2	1.8
$t+1$	0	0	0	0	.8	.9	.6	.4	4.9	9.1	2.4	1.6
$t+2$	0	0	0	0	.5	.7	.4	.3	4.0	7.7	2.1	1.8
$t+3$	0	0	0	0	.3	.5	.4	.5	3.5	5.9	2.1	1.8
Rule: $\Delta M = 2.0$ per quarter after first quarter												
t	-114	-211	-148	-84	1.0	1.0	1.0	1.0	4.7	9.1	3.2	1.8
$t+1$	+70	+161	+16	-17	0.2	.1	.5	.6	2.0	2.0	2.0	2.0
$t+2$	-3	-31	-4	+4	0.2	.2	.4	.3	2.0	2.0	2.0	2.0
$t+3$	-1	+3	+3	+3	0.2	.1	.3	.5	2.0	2.0	2.0	2.0

* Definitions of variables are given in footnotes to Tables 1, 2, 3, and 4.

† Model 1=FRS-MIT, Model 2=FRS-MIT-Chicago, Model 3=Brookings, Model 4=Ando-Goldfeld.

SOURCE: All entries were calculated from basic data in Tables 1, 2, 3, and 4, under the assumptions of policies 1, 2, and 3.

more false information than the interest rate. In general, any policy that leads to sufficiently dampened oscillations of ΔB around zero will favor interest rates as an indicator. The examples presented here were chosen because they are frequently mentioned policies.

A second important determinant of current changes in r and M is a changing rate of current government expenditures. As with past changes in B , ranges of fiscal shocks ΔF can be derived which make r and M true and false indicators of the effect of current ΔB on income. Choosing the monetary policy $\Delta B = \$1.0$ billion per quarter as an initial condition assures that the separate effect on B on Y each quarter is expansionary. In the first quarter of this experiment, FRS-MIT implies the ranges summarized below.

FIRST QUARTER				
	r true M false	r true M true	r false M true	ΔF billion dollars
FRS-MIT	-23.5	0	21.7	

To the extent that changes in government spending are not likely to exceed \$20.0 billion in either direction in a single quarter, FRS-MIT implies that it is unlikely that the quality of either r or M as indicators will be affected by fiscal policy in the first quarter.

FRS-CHICAGO-MIT implies slightly different ranges as reported below.

FIRST QUARTER				
	r true M false	r true M true	r false M true	ΔF billion dollars
FRS-CHICAGO-MIT	-65.0	0	35.0	

As with FRS-MIT, it appears unlikely that fiscal policy changes would be large enough to affect the indicator quality of either r or M in the first quarter of this monetary policy.

Brookings implies that ΔF is never likely to affect the indicator quality of either r or M . For monetary policy $\Delta B = \$1.0$ billion, Brookings requires ΔF of roughly (plus or minus) \$100 billion

to make either r or M yield false information in either the first or second quarter of the experiment. Ando and Goldfeld imply that fiscal policy can never affect the quality of M as an indicator of the effects of B on Y because ΔF causes no change in M (see Table 2). For Ando and Goldfeld, $\Delta F > \$465$ billion in the first quarter, or $\Delta F > \$31$ billion in the second quarter cause r to yield false indicator information.

A strong, though not surprising, pattern emerges from these results. All models imply that when B and F are moving in the same direction, M , interpreted by the direction of change criteria, will always yield correct information about the thrust of current monetary policy actions or income. Changes in r are more or less likely to yield false indicator information depending on the relative sizes of changes in B and F when both B and F move in the same direction. If B and F move in opposite directions, r will always yield correct information while changes in M will yield false information if ΔF is sufficiently large relative to ΔB . Thus, to the extent that B and F tend to move in the same direction, r may and M cannot give false signals about the thrust of monetary policy on income. If B and F move in opposite directions, M may and r cannot give false signals.

III. Conclusions

All four econometric models considered were found to imply that two sources of effects on current interest rates and money stocks other than current monetary policy actions could, under certain conditions, cause either interest rates or money stocks to yield false information about the direction of effect of current monetary policy actions on income. These two sources of effects are (1) past monetary policy actions as summarized by changes in the amount of base money or reserves and (2) current fiscal policy actions represented by changes in government expenditures. No other predetermined shocks were considered.

The following general statements are implied by all four of the models. (1) Past monetary policy actions (ΔB) make changes in the money stock yield false indicator information only if base money changes direction, and if the absolute change in base money decreases sufficiently from its last quarter level. (2) Past monetary policy actions make changes in the interest rate yield false information only if base money moves in the same direction in succeeding quarters and decelerates sufficiently. (3) Current changes in government expenditures (ΔF) can make money yield false indicator information only if ΔF and ΔB have opposite signs and ΔF is very large relative to ΔB .

(4) Current changes in government expenditures can make the interest rate yield false indicator information only if ΔB and ΔF move in the same direction and ΔF is very large relative to ΔB .

Considering in a slightly different light the states of nature necessary to make one indicator candidate yield true and the other false information may clarify the implications of the models. In the first state base money increases every quarter and accelerates and decelerates from time to time, while government expenditures always increase and on occasion increase rapidly relative to the change in base money. The models imply that the interest rate is more likely than money to yield false indicator information in this world due to past monetary policy actions or current fiscal policy actions. The money stock is also affected by these predetermined shocks, but nonetheless continues to yield correct information about the thrust of current policy. In the second state of nature the monetary base changes direction every quarter, the absolute values of changes in B tend to increase and decrease over time, while government expenditures always change in the opposite direction of base money and the absolute values of these changes are occasionally large relative to changes in base money. In this world the interest rate is a better indicator than the money stock according to these models.

When only the effects of past monetary actions and current fiscal actions are considered, these models imply the money stock is a more reliable indicator of the impact on income of current policy actions than is the interest rate. This follows from the unlikely variability of monetary and fiscal policy necessary to make the interest rate a better indicator than the money stock. We would have to observe frequent reversals in the direction of change in base money or reserves, erratic magnitudes of these changes, and monetary actions affecting income in the opposite direction from fiscal policy in order to reverse this conclusion.

Earlier it was asserted that neither the income-expenditure theory nor the quantity theory implies a position on the indicators issue. It now has been shown that all four of these models, which are based on the income-expenditure theory with an explicit interest rate transmission mechanism, imply that a wide range of behavior of past monetary policy and of current fiscal policy instruments makes money a true and the interest rate a false indicator of monetary policy. It has thus been demonstrated by counterexample that a basic theoretical position does not imply a position on the indicators issue.

REFERENCES

1. ALBERT ANDO AND STEPHEN GOLDFELD, "An Econometric Model for Stabilization Policies," paper presented to the Committee on Financial Analysis, Nov., 1966.
2. KARL BRUNNER AND ALLAN H. MELTZER, "Indicators and Targets in Monetary Policy," paper presented to the Univ. of California at Los Angeles Conference on Targets and Indicators, forthcoming publication.
3. ———, "The Meaning of Monetary Indicators," in *Monetary Process and Policy: A Symposium*, ed., George Horwich (Richard D. Irwin, Inc., 1967).
4. RICHARD G. DAVIS, "The Role of the Money Supply in Business Cycles," *Monthly Review*, Fed. Res. Bank of New York, Apr., 1968.
5. FRANK DE LEEUW AND EDWARD GRAMLICH, "The Federal Reserve-MIT Econometric Model," *Fed. Res. Bul.*, Jan., 1968.
6. MILTON FRIEDMAN, "The Role of Monetary Policy," *A.E.R.*, Mar., 1968.
7. ———, *The Optimum Quantity of Money and Other Essays* (Aldine Pub. Co., 1969).
8. GARY FROMM AND PAUL TAUBMAN, *Policy Simulations with an Econometric Model* (Brookings, 1968).
9. ROBERT J. GORDON, "The Brookings Model in Action: A Review Article" (unpublished manuscript, 1969).
10. LYLE GRAMLEY AND SAMUEL CHASE, "Time Deposits in Monetary Analysis," *Fed. Res. Review*, Oct., 1965.
11. JOHN G. GURLEY AND EDWARD S. SHAW, *Money in a Theory of Finance* (Brookings Institution, 1960).
12. PATRIC H. HENDERSHOTT, *The Neutralized Money Stock* (Richard D. Irwin, Inc., 1968).
13. GEORGE G. KAUFMAN AND ROBERT D. LAURENT, "Three Experiments with Simulating a Modified FRB-MIT Model," paper presented to the Federal Reserve System Committee on Financial Analysis, 1969.
14. ROBERT A. MUNDELL, *International Economics* (Macmillan Co., 1968).
15. JAMES TOBIN AND WILLIAM BRAINARD, "Financial Intermediaries and the Effectiveness of Monetary Controls," *A.E.R.*, May, 1963.
16. T. SAVING, "Monetary Policy Targets and Indicators," *J.P.E.*, Aug., 1967.

DISCUSSION

PETER A. FROST: The chairman asked me to concentrate my remarks on Michael Hamburger's paper. The absence of comments on the other papers reflects the space constraint and not my view of their quality or relevance.

Hamburger presents empirical results relating to, one, how sensitive various monetary aggregates are to changes in endogenous variables and, two, the extent to which various monetary aggregates affect GNP. The second question covers the familiar ground of the Andersen-Jordan study and is important for determining an appropriate monetary indicator, since everyone seems to agree that the indicator should have a strong and predictable influence on economic activity. There seems to be much less agreement about whether the monetary indicator must be an exogenous variable. I would answer in the negative, since a monetary indicator does not have to be exogenous to accurately indicate the effect of monetary policy on economic activity. However, if the monetary indicator is not exogenous, care must be taken in estimating the effect that changes in the indicator have on GNP. Most of the discussion about whether various monetary aggregates are exogenous has centered on this estimation problem.

Unfortunately, the statistical tests run by Hamburger cannot determine if a variable is exogenous. The question of whether a variable is exogenous is inseparable from the question of which model predicts best. As an example of why simple correlations cannot determine if a variable is exogenous, consider the following system of equations.

$$X_1 = a_0 + a_1X_2 + a_2X_3 + u_1$$

$$X_2 = b_0 + b_1X_1 + b_2X_4 + u_2$$

In this model, X_1 and X_2 are endogenous variables and X_3 and X_4 are exogenous. If a_1 and b_1 are statistically significant, but of opposite signs, the simple correlation between X_1 and X_2 could easily be zero. An example of this is the lack of correlation between price and quantity when both the supply and demand curves have been shifting. Also, a significant correlation between X_1 and X_3 or between X_2 and X_4 does not indicate that either X_3 or X_4 are endogenous variables. If this model predicts better than any other model, then X_1 and X_2 can be considered endogenous and X_3 and X_4 exogenous.

Hamburger notes that most observers agree that, during a large part of the post-World War II period, the Federal Reserve concentrated on either free reserves or some money market interest rate and let the money supply passively adjust to changes in the public's demand for money. Critics of the Andersen-Jordan study have argued that this behavior makes the money supply an endogenous variable and raises serious questions about the direction of causation.

The two-stage regressions reported below constitute a rudimentary attempt to find out the degree to which the parameter estimates are biased by ignoring this type of a feedback from GNP to money. Two basic

assumptions are made. First, it is assumed that the Federal Reserve controls some money market rate by buying or selling securities in the open market whenever the rate deviates from the desired level. This makes the money supply a function of the desired interest rate and the public's demand for money. Second, a standard Keynesian demand for money function is assumed. This means that the money supply will vary directly with GNP as long as the Federal Reserve maintains a given interest rate. The following two equations are estimated.

$$Y = a_0 + a_1M + a_2G + u_1$$

$$M = b_0 + b_1FRP + b_2Y + u_2$$

The first equation is the reduced form of a system of equations in which the money supply is treated as an exogenous variable. Y is the quarterly percentage change in GNP, M is the quarterly percentage change in the narrowly defined money supply, and G is the quarterly percentage change in high employment expenditures. The second equation is the money supply equation resulting from the assumptions mentioned above. FRP is the Federal Reserve policy variable. It is the quarterly sum of the Brunner-Meltzer index of the directives of the Federal Open Market Committee. Brunner and Meltzer assigned a value ranging from +1 to -1 to each FOMC directive depending on the degree to which it called for easing or tightening money market conditions. A value of zero is assigned whenever the current conditions are to be maintained.

The parameter estimates for the above equations are shown in Table 1. The estimates of a_1 increase by a factor of from 3 to 6 times and have higher t -values when two-stage least square (TSLS) estimates are used instead of one-stage least square (OLS) estimates. This result is contrary to the expectations of the Andersen-Jordan critics. It is interesting that the greatest improvement comes in the period from 1952 to 1960, the period during which the St. Louis equation does the worst. While it is clearly dangerous to make any inferences about lags from regressions that exclude lagged terms, the large increase in the value of a_1 when TSLS is used suggests that the lag in response to changes in the money supply may be smaller than OLS estimates would lead us to believe. Only one of the money supply regressions is reported because of the relatively poor fit.

In the last part of his paper, Hamburger regresses GNP on various lagged monetary aggregates to determine which one has the greatest effect on GNP. He argues that his results challenge the monetarists position because the R -squareds for the bank credit regressions are higher than the R -squareds for the money supply regressions. This is indeed an interesting result, but it certainly is not sufficient to reject the monetarists position. The results presented above show the extent to which the parameters can change when the money supply is treated as an endogenous variable. Since the demand for money is usually viewed as being a function

TABLE 1
GNP Regressions

Year	Type	Int.	<i>M</i>	<i>G</i>	<i>R</i> ² / <i>DW</i>
1952-68	OSLS	.848 (4.6)	.768 (4.1)	.049 (.9)	.208 1.31
1952-68	TSLS	-.20 (-.7)	2.49 (6.0)	-.035 (-.7)	.361 1.72
1952-60	OSLS	.696 (2.7)	.993 (2.7)	.031 (.4)	.140 1.21
1952-60	TSLS	-.08 (-.2)	3.02 (4.4)	-.024 (-.3)	.342 1.80
1961-68	OSLS	1.31 (5.2)	.410 (2.3)	.039 (.6)	.105 1.68
1961-68	TSLS	-.68 (-.8)	2.36 (2.9)	.066 (1.1)	.174 1.64

Money Supply Regressions

Year	Type	Int.	<i>FRP</i>	<i>Y</i>	<i>R</i> ² / <i>DW</i>
1952-68	OSLS	.376 (2.8)	-.232 (-1.4)	.212 (2.7)	.221 .85
1952-68	TSLS	-.120 (-.2)	.190 (.4)	.556 (1.6)	.167 .90

of wealth or permanent income, it would not be surprising to find that the public's demand for bank credit is more closely related to GNP and, consequently, that the reverse causation is greater in the bank credit regressions. Unfortunately, one cannot use the *R*-squareds from TSLS regressions like those reported above to determine the relative importance of money and bank credit. This question will have to be settled by exploring the implications of more detailed models of the monetary mechanism.

LEONALL C. ANDERSEN: Holbrook and Shapiro in their paper have attacked a problem which is of great concern to those of us who are actively engaged in monetary management. Recently there has been considerable controversy over choice of intermediate targets for Federal Reserve short-run monetary actions. Market interest rates have been the most frequently used guide for quite a few years, but in recent years there has been growing support for using some monetary aggregate such as the money stock, the monetary base, total reserves, or bank credit as an intermediate target. Holbrook and Shapiro have examined the relative merits of market interest rates and the first two aggregates I just mentioned.

In developing a model for their purposes, they proceed along familiar *IS-LM* lines. Many economists have argued the relative merits of interest rates and the money stock as targets for Federal Reserve economic stabilization actions by pointing out the importance

of the comparative stability of the *IS* and *LM* relationships. If the major source of economic instability stems from the real sector (*IS*), the money stock is a preferred target. On the other hand, if the demand for money is a major source of instability (*LM*), the interest rate is to be preferred. Most of the discussions of these points have been mainly theoretical and, as a result, they have been interesting but not too useful for structuring Federal Reserve operations. By moving forward to empirical analysis, Holbrook and Shapiro have made a useful contribution to the current debate over selection of an intermediate economic target.

Although their results—that the narrowly defined money stock is a better intermediate target than the interest rate—support the general conclusions of my own research, I have some reservations, as they also have, regarding their work. My reservations are of two kinds: first, their acceptance of the usual Keynesian assumption of a constant price level and, second, their use of a crude quadratic loss function.

The authors' principal conclusion is that "optimal policies depend not only on the structure of the economy, but on the particular nature of the uncertainties facing the policy-maker." Generally speaking, this is the same as Theil concluded over a decade ago. The model the authors use has at least two serious deficiencies: it is linear but the real world is not and the price level is assumed constant. The latter implies that the central bank can control real balances by controlling nominal money balances and thus pursue intermediate interest

rate targets more or less directly. If prices were a variable, the reduced form containing the interest rate would have many more parameters to forecast providing more "uncertainties," regardless of the relative stability of the *IS* versus the *LM* curves. Further, unless prices were determined via "wage-push," the system would become nonlinear, with the result that the errors would be multiplicative rather than additive.

Using a crude quadratic loss (preference or disutility) function is subject to a variety of standard criticisms. A level of real GNP above the target should not be symmetrical to a level of real GNP below the target, as implied by the squared deviation term in the loss function. This is the "asymmetrical preferences" argument commonly directed toward the unemployment rate. Further, the target value of real GNP cannot be chosen arbitrarily, as the authors have done, but it must be on the production possibility frontier; the loss function must first be minimized subject to a potential GNP constraint in order to compute the value of the target.

The authors have actually picked a fairly restricted case to study; namely, the one target—one instrument (actually, intermediate target) case with a single period horizon. All that keeps the use of any one of the intermediate targets from hitting the one ultimate target continually is the variance in the ultimate target associated with each intermediate target (and target variance enters the loss function). Their paper sheds no light on the problem of policy mixes, multiple targets, the evaluation of alternative time paths of response, and the optimal timing of policy actions.

I now turn to a few comments on the other two papers. Richard Zecher's paper is a useful contribution to the indicator issue. I find it particularly interesting that, in the four major large-scale econometric models examined by Zecher, circumstances which are most likely to occur lead to the money stock being a better indicator of the thrust of monetary actions than market interest rates. His paper, however, has a problem similar to the Holbrook-Shapiro study. None of the four models allow for the influence of expected movements in the price level on market interest rates. Zecher presents a good discussion of this problem, but he is restricted from taking it into consideration by the nature of the models from which his conclusions are drawn.

Michael Hamburger examines in considerable detail the exogenous-endogenous controversy regarding single equation estimates of the parameters relating changes in GNP to various monetary aggregates. Many have contended that the money stock is an endogenous variable and, therefore, parameters estimated by ordinary least squares (OLS) of the response of GNP to changes in the money stock are biased. It is further contended that the bias is so great as to render such regressions useless in conducting monetary policy. Hamburger has attempted to answer these criticisms, but I am afraid that the critics will not be convinced by neither his evidence nor by evidence I have presented. I also find the search of many for the "most exogenous" monetary variable to be particularly unrewarding.

We should keep the question of simultaneous equation bias in its proper perspective. It is strictly a statistical property of estimators, and there is no accepted

test for its presence. Statistical theory at the present time can tell us little about this problem in economic research. Formal discussions of bias are based on the asymptotic properties of large size samples and little is known about bias in the limited size, finite samples available to economists. Monte Carlo methods have been used to measure bias in small finite samples, but conclusions based on such studies do not apply to regressions which include lagged, endogenous variables.

It is well known in statistics that parameters estimated by OLS have smaller variances than those estimated by such other procedures as two-stage least squares. Thus, in selecting estimation procedures, one must trade-off bias against variance of parameter estimates. In view of the unsettled nature of these issues in economic research, there is no clear-cut case for asserting that there is an obvious bias in the OLS parameter estimates of such a magnitude as to more than offset the gains of smaller variances of parameter estimates, thereby making the results highly unreliable.

In conclusion, I want to compliment the authors of all three papers for making valuable contributions to the indicator and target controversy. They have moved this controversy from strictly theoretical discussions to empirical evidence. Its latter form will be particularly useful to those charged with the responsibility for monetary policy.

THOMAS J. SARGENT: A monetary indicator is a variable the value of which tells whether monetary policy is becoming more or less restrictive, given the influences of current and past fiscal policy, past monetary policy, and the stochastic elements affecting investment, saving, and portfolio balance. Presumably, advocates of the interest rate as an indicator interpret a fall in the rate as a symptom of a less restrictive monetary policy, while advocates of the money supply interpret a larger money supply or larger rate of change of the money supply as indicative of a more expansive monetary policy. However, it appears that the problem of choosing an indicator has no content over and above the problem of choosing the optimal monetary instrument. Particularly when viewed in the light of Holbrook and Shapiro's paper, Zecher's paper illustrates the fact that the information needed to assess the appropriateness of alternative monetary indicators is exactly the information required to choose between alternative monetary instruments. Moreover, what is required in order for the money supply (or interest rate) to fulfill the indicator function adequately is the fulfillment of precisely those conditions which will make the money supply (or interest rate) a superior monetary instrument.

A glance at the major points of controversy in the discussion of monetary indicators makes it clear that there is really no separate indicator problem. Several limitations of the use of the interest rate on government bonds as an indicator of monetary restraint have been cited in the literature. What these limitations have in common is that they are all things which make aggregate demand an unstable function of the current nominal interest rate on government bonds. Thus, it has been noted that the aggregate demand schedule shifts in response to past accelerations in output, past changes in monetary aggregates, and current and past fiscal policy. It has been

repeatedly emphasized that the schedule will also shift as the anticipated rate of inflation changes, since anticipated real rates of return are presumably the relevant ones for spending and saving decisions. Moreover, even ignoring the effects of anticipations of price changes, the interest rate on government bonds may not always be a reliable proxy for the cost of capital facing firms and individuals. This is another reason that aggregate demand will be an unstable function of the current nominal interest rate on government bonds. The instability of that schedule is what makes the nominal rate on government bonds a poor indicator of monetary restraint, since it implies that a decrease in the nominal bond rate cannot be attributed to a movement along the schedule. Anything which makes that schedule unstable qualifies as a factor making the interest rate on government bonds a poor indicator.

Hence the entire discussion of the interest rate as an indicator can be summarized succinctly by saying that the appropriateness of the nominal rate on government bonds as an indicator depends on the stability of the *IS* curve. On the other hand, the conditions which must be met for the money supply (or the monetary base) to be an appropriate indicator involve the stability of the *LM* curve. Thus, if the demand function for money often shifts, say as a result of the monetary authorities varying the maturity structure of the government debt in the hands of the public, the money supply or the base becomes a weak indicator. The base will be a weaker indicator the more use the authorities make of alternative instruments such as discount rate or reserve requirement changes; i.e., the more unstable is the relationship between the money supply and the base. Generally, any forces which increase the stochastic components of the demand and supply functions for money will diminish the usefulness of the money supply or base as an indicator.

The fact that the resolution of the indicator problem hinges on the relative stability of the *IS* and *LM* curves establishes the identity between the indicator problem and the choice-of-instruments problem; for as Bailey and Kareken have pointed out, the relative stability of the *IS* and *LM* curves is what determines whether money or the nominal interest rate is a superior instrument variable. If the *LM* curve is more unstable, the monetary authorities can minimize fluctuations in income by using an interest rate instrument. By doing so, they can offset the effects on output of the disturbances affecting the *LM* curve. On the other hand, if the *IS* curve is relatively less stable than the *LM* curve, fluctuations in output will be reduced by using the money supply as the instrument. Of course, if it does turn out that the money supply is the optimal instrument variable, movements in the money supply will also be an appropriate indicator of movements in the direction of monetary policy. Similarly the interest rate will be an adequate indicator if and only if it is also the optimal monetary instrument. The upshot is that there is not a separate indicator problem over and above the problem of choosing an optimal monetary instrument.

This is brought out quite well by the experiments Zecher has performed. If Zecher finds in a given model that the interest rate fails to vary inversely with the

monetary base while the money supply generally varies directly, it means that the *IS* curve is relatively unstable while the *LM* curve is more stable, implying that the money supply is a superior instrument. Zecher's simulation experiments should thus be viewed as providing information about the same problem discussed by Holbrook and Shapiro. The advantage of Zecher's procedure of examining large econometric models is that the models hopefully catch more of the critical elements involved in choosing the optimal instrument. For example, at least one of the models studied by Zecher incorporates some of the effect of the money supply on prices and the effect of anticipated price changes on yields. Those effects, which are omitted from the Holbrook-Shapiro model, add importantly to the instability of the *IS* curve relative to the *LM* curve and are important elements in the arguments in favor of the money supply as an instrument. The monetarists argue that a disturbance of the *IS* curve is aggravated if the monetary authorities pursue a target nominal interest rate, since the money supply changes required to achieve the target will cause price changes and changes in anticipated inflation which in turn push the *IS* curve even farther in the direction of the initial disturbance. Recently the monetarists have been placing much stress on this argument. (It is interesting that Holbrook and Shapiro find that the money supply is a superior instrument even while using a model which fails to incorporate this argument.)

In assessing the usefulness of the experiments Zecher has performed, it is necessary to determine whether he has given proper attention to all of the important sources of shocks which may make the *IS* and *LM* curves unstable. By incorporating shocks which affect only fiscal policy and the lagged monetary base, Zecher has focused on forces which make the *IS* curve unstable; however, he has neglected shocks, say arising from changes in the maturity structure of the public debt or changes in Regulation Q, which would destabilize the *LM* curve. More importantly, simulations of the kind Zecher has performed are of limited usefulness in deriving the implications of large econometric models for the optimal choice of monetary instrument. At least up to a linear approximation, the variances of output and any other goal variables can be calculated analytically from the estimated parameters of the models. Alternatively, if a linear approximation seems too crude, Monte Carlo techniques can be used to estimate the distributions of goal variables and to determine their dependencies on various parameters. By using one of these techniques, it is possible to make use of the covariance matrix of the residuals, which contains some of the essential empirical data required to choose a monetary instrument optimally. Zecher's simulations make no use of these important data.

Zecher's conclusion that a view about the relative potency of monetary and fiscal policy does not necessarily imply anything about the proper resolution of the monetary instruments problem is certainly correct. To show this it is sufficient to note that relative potency depends on the relative slopes of the *IS* and *LM* curves, while resolution of the instruments question depends on their relative stability as well as on their slopes.

PROGRAMMING MODELS OF NATIONAL ECONOMIES

PLANNING WITH FACTS: THE CASE OF KOREA

By ROGER D. NORTON

International Bank for Reconstruction and Development

I. Introduction

Korean economic planning has been more a process of experimentation than the orderly, repetitive sequence of calculations usually associated with the term planning. This was almost inevitable in view of the conditions in the South Korean economy in recent years. The rate of economic expansion has accelerated at a pace impossible to foresee, and significant policy reforms have been undertaken in several areas. As a consequence, the planners' assessments of the constraints to development—and to some extent the goals of economic policy—have changed.

To briefly indicate the pace of expansion, the GNP growth rate has been about 11 percent per year, in real terms, for the past seven years, and commodity exports have expanded by 40 percent per year for the past twelve years. The most significant policy reforms were enacted in the four years 1964–67. These included an exchange rate devaluation of about 100 percent, a doubling of commercial bank interest rates,¹ a restructuring of the tariff system on the principle of effective protection rates, and a revision of the tax structure.

At the time the Second Five-Year Economic Development Plan (1967–71) was prepared, the economic environment was strongly colored by the poor performance of the first postwar decade. One of the first major tasks of the Plan was to determine the highest feasible growth rate—the size of the Plan. The major obstacles appeared to be insufficient domestic savings and inadequate export earnings. Aggregate and sectoral consistency tests were constructed and, as the Plan outlines emerged, it appeared as though an annual real growth rate somewhat in excess of 8 percent was possible. This was a bold conclusion at the time, and recent history dictated a cautious approach, so the number 7 percent was adopted and emphasis was placed on exceeding that if possible.

II. Design of the Plan

Korean planning has involved rather detailed analysis of public options in some areas and yet

has touched only lightly on other sets of problems. The second Plan itself was very much oriented toward investment programs in industry and social overhead, by sector and project, and the savings and export problems. Employment, education, agriculture, and social welfare received much less attention.

The broad aims of the planning process in Korea may be described as, first, articulation of the society's economic goals; second, calculation of a feasible growth path which goes far toward satisfying these goals; third, identification of major constraints and ways to alleviate them; and, fourth, construction of detailed investment programs and public policies to move the economy along the preferred path. This is to be a continuous process; explicit provision was made in the Second Plan for a series of annual plans, to translate the five-year plan into more specific components and to ensure review and revision of the overall plan design. The relationship of these annual plans, or Overall Resources Budgets (ORB's), to the actual budgeting process was ambiguous in the Second Plan document. In the subsequent attempts to define that relationship, several institutional forms of the ORB have been tried. Long-term perspectives also were incorporated in the Plan, in the form of fifteen-year projections which defined the long-term strategies.

For the short- and medium-term analysis, the planning process has been multilevel, in the sense that consistency was required among the programs and policies at the aggregate, sectoral and project levels [2]. Short of using formal, multilevel optimization, considerable effort was put into achieving consistent plans at these three levels, particularly in the realm of investment programs and financial policies.

The major quantitative tools for the Second Plan were an input-output consistency model [2], a macroeconomic model, and rate-of-return calculations for projects. To insure consistency of the annual investment plans with reasonable price stability, a short-term stabilization model was subsequently developed,² and to insure con-

¹ So that the real rate of interest rose to about 20 percent on long-term deposits.

² For discussions of this model at two phases in its evolution, see [6] [7].

sistency of demand and supply for unusually lumpy investments, the results of a mixed integer programming model were used [17]. Finally, to define goals and feasible paths for regional development programs, a regional linear programming model was developed [11].

III. Goals

Rapid aggregate growth was the dominant goal of the Second Plan and the annual Overall Resources Budgets. Growth is not presented as a goal in itself, but rather as the best means for Korea to achieve more control over its own economic destiny, to attain better living standards for future generations, and to reduce unemployment. For the government of President Chung Hee Park the first of these three aims became an important political commitment. In the President's own words prefacing the Second Plan, "We are determined to achieve self-sufficiency in the foreseeable future."³

The phrase "self-sufficiency" does not mean economic autarky, but rather elimination of dependence on foreign assistance. It takes on added meaning in Korea, which for centuries was dominated by China or Japan and now finds itself with little room to maneuver between the U. S. and Japan, and Russia and China.

A high population density⁴ and a generally unindustrialized economy have produced a large number of unemployed and underemployed. For the medium term, a growth-oriented policy was seen as the best employment policy. The relatively high literacy rate (about 90 percent) and the cultural stress on hard work provided a readily employable industrial labor pool. However, during the Plan formulation, employment criteria generally were applied to projects and sectoral programs when the growth criteria were indecisive [10], so reduction of unemployment should be regarded as one of the major five-year planning goals, albeit secondary in importance to growth.

The plan also refers to concern over the income distribution, regional and otherwise, but the actual program and project design showed little serious concern for these issues. Relative prosperity has brought increasing public concern over income distribution, and it has been announced that the Third Plan will devote more attention to equity issues and will include a set of regional development plans.

Price stability is mentioned not at all in the discussion of the basic five-year objectives, nor in the

list of subsidiary targets. This omission is partly explained by the fact that income increases are stated in real terms and partly by the emphasis on price stability in the fifteen-year outlook chapter, but nonetheless it is remarkable in view of Korea's experience with hyperinflation in the late 1940's and with rapid inflation immediately after the Korean war and again in 1963-64.

This omission may be interpreted in the following ways: the government felt that inflation at the current rate of 8-10 percent per year would be tolerable; and the government felt confident of its ability to restrain the price increases to that range.

IV. Strategies

The fifteen-year perspective clearly stated the four key elements of long-run planning strategy in Korea: export expansion, domestic capital mobilization, efficient manpower and technology utilization, and continuing financial stability.

For the medium term, the input-output model exercises had confirmed the importance of exports and domestic savings as limits to growth. Korean planners were convinced that the only bottleneck to rapid export expansion lay on the supply side, and thus to fulfill both the export and employment/technology strategies, investment planning became important. To assure financial stability and to make a closer determination of the savings constraint, financial planning of a sort was adopted. These two concerns—investment allocation and financial intermediation—became the analytic focus of the first three annual Overall Resources Budgets under the Second Plan.

The Economic Planning Board is legally charged with the responsibility for planning the allocation of public sector investment and foreign capital inflows. Since some foreign capital inflows are tied to specific projects, planning in this area became "indicative" to some degree.⁵ The plan documents have carried lists of projects for which financing was sought and also lists of expected sectoral demands for new plant capacity.

At the beginning of the plan formulation process, there were very few viable investment project proposals on hand, and a substantial effort was required to develop each worthwhile proposal. Therefore the input-output model was used to indicate priority sectors for project formulation.⁶

³ Indicative planning is usually associated with the French plans; see, for example [3].

⁶ In the words of the Second Plan: "It is recognized that the process of generation of project proposals is a key part of the formulation of an investment program. The composition of the final investment program can be

³ Second Five-Year Plan [12], introductory statement.

⁴ 305 persons/sq. km.

Sector-by-sector assessments were given, as illustrated by the following excerpts:

Paper products. Investment in this sector will be increased considerably in the future. There is an urgent need for new investment in addition to the projects already in progress . . .

Nonferrous metals. There are sufficient project proposals in smelting and refining of nonferrous metal ores, but more proposals are needed for processing facilities such as extrusion and rolling plants.⁷

These kinds of statements were drawn from careful examination and revision of input-output model projections and tabulation of on-going major investments. To implement the Plan sectoral investment projections at the 43-sector level, and to disaggregate these into projections for about 150 manufacturing sectors, required unusual efforts to gather new technical data [2]. A "man-machine" iteration was established, in the words of Professor Adelman, whereby industrial experts contributed to revision of the technical coefficients in the model, and the model projections of interindustry flows contributed to revision of the experts' estimates of future demand. For the continuing annual ORB work on investment guidelines, further cooperation was elicited from the Industry Committees and from the commercial banks.

Clearly the attempt to keep track of the literally hundreds of investment projects in the larger manufacturing plants, a necessary exercise for annual projection of further sectoral capital requirements drew heavily upon the planning bureau's manpower. Partly for this reason,⁸ the indicative planning exercise was dropped in the 1969 and 1970 ORB's, and attention was confined to the public sector investment program and aggregate magnitudes.

While the sectoral investment programs were enunciated in pursuit of consistency on the inter-

industry demand side, flow-of-funds projections were made for the supply side (savings) consistency tests.

The working hypothesis was that identity of savings and investment in real, aggregate terms is no guarantee that the investment programs could be financed in a noninflationary manner. Savings take various forms, and each form is appropriate only for certain types of investment. To some extent the government and private sectors compete for claims on financial assets. This competition is reflected partly in the distribution of financial intermediation chores between the central bank and the commercial banks [6]. Furthermore:

Since the capital market is relatively underdeveloped other techniques of savings intermediation (perhaps aside from fiscal means) are extremely imperfect. It has consequently brought about a heavy burden for domestic financing of major investments upon the banking sector. . . . This situation implies a stronger interdependency between monetary controls and investment programs in less developed countries than in many advanced countries [7].

In the 1968 ORB, alternative flow-of-funds projections for the year were presented for different assumptions about monetary policies and exogenous variables. The inflationary implications of each program were explained in accompanying documents.

The basic tool for these studies was the stabilization model—a set of estimated equations linking the financial and real sectors. Portfolio choice theory provided the basic form of the equations; desired asset holdings of each type depend on expected relative returns. There are sets of equations explaining desired or expected corporate asset positions, desired corporate liabilities, banking system liabilities, bank assets, and price formation. Prices were explained in the model by the discrepancy between supply and demand for liquid assets, with lags.⁹

Some of the conclusions of the first stabilization model studies were as follows [6]:

1. A study of disaggregated flows of savings is required for a proper feasibility test of the annual growth and investment targets. The real demand for liquid assets, plus banks' retained earnings and government loans, impose a ceiling on banks' portfolio investment if inflation is to be maintained within the desired limits.

influenced heavily by the types of proposals submitted to the planning agencies. Therefore, the Government has assumed a guidance function in disseminating information on the types of industries which will be facing heavy demands for capacity expansion over the next five years. This is in effect the first stage in the long process of project selection, and it serves to channel project proposals into the areas in which they are most needed" (p. 80).

⁷ *Overall Resources Budget 1968 (First Draft)*, Economic Planning Board, Seoul, Mar., 1967. For the 1968 ORB, two drafts were published; the first, in the spring, to serve the indicative planning function, and the second, in the fall, to conform to the actual budget.

⁸ And also because the modified 1965 I-O tables were becoming more out of date with the 15–20 percent expansion registered each year by industry.

⁹ This hypothesis, derived from an explanation of the public's desired portfolio structure, was first applied to Korea by John Gurley, Hugh Patrick, and E. S. Shaw [4].

2. The excess of money supply over real demand for money, current and lagged, appears to provide an explanation of changes in the general price level. In contrast to the situation in more developed countries, the real demand for money in Korea is fairly unstable over time.

3. Price expectations influence inflation; the substitution between money and short-term savings deposits is significant empirically, and ignoring this effect overstates the real demand for money, and thus understates the inflationary pressures, when savings deposits are growing rapidly.

4. Lagged supply of and demand for money, lagged by at least six months, is more significant than current values in determining current rates of price increases.

One by-product of these discussions was a change in the official definition of the money supply, an indicator closely watched by guardians of the price level. Late in 1967, the definition was expanded to embrace short-term time deposits.

As with the other planning exercises, the stabilization model required development of new data. The Bank of Korea had published annual flow-of-funds tables since 1963. To estimate the equations, it was necessary to develop a longer time series; with previously unpublished data, the Bank of Korea was able to construct additional flow-of-funds tables for 1958-62.

The financial model projections were continued for the 1969 ORB but dropped for the 1970 ORB as attention shifted to preparations for the Third Plan. In a longer-run perspective, the main contribution of these exercises probably has been the fostering of a sharpened awareness of financial constraints. Future work along these lines probably will employ a quarterly econometric model. The Bank of Korea recently has developed quarterly estimates of national accounts in response to the need for better tools for short-run policies.

Perhaps as an acknowledgement of the growing public concern over disparities in regional rates of industrialization, the Second Plan also called for the formulation of regional development plans.

After the Plan was issued, the Economic Planning Board proceeded to draft a lengthy preamble to regional planning which set out administrative measures and research projects necessary to complete meaningful and coordinated regional plans for inclusion in the Third Plan [5]. Research projects on interregional input-output tables, interregional transportation costs, and interregional flows of funds were financed by the Economic Planning Board and the Ministry of National Construction. The Bank of Korea again contrib-

uted by issuing regional national accounts estimates beginning from 1965.

The first concrete accomplishment of the interministerial regional planning effort was the definition of regional boundaries for planning purposes—a task which consumed more than a year. The ultimate geographical partition was similar to the French planning partition in the sense that regions were defined at different levels of aggregation: at one level, four supraprovincial regions were defined; at another level, the nine provinces serve as a set of regions; and at yet another level, there are seventeen regions.

Due to the complexity of the regional planning problem, foreign assistance on a larger scale has been requested. A United Nations Special Fund team has been engaged by the Ministry of National Construction to help draft the various facets of regional plans. The regional input-output tables are being used in a regional investment allocation model.

At this point, the regional planning still is in an embryonic stage, but it may be said that institutional arrangements are more critical to a successful regional planning effort than to other forms of planning, simply because so many different public agencies and private groups are affected.

A subsidiary element of the Plan strategy was more efficient allocation of foreign resources via trade policies. As a result of the availability of the input-output tables, the Ministry of Finance was able to authorize a study for the measurement of effective rates of protection at the 270-sector level of disaggregation, in preparation for the tariff revision late in 1967.¹⁰

The I-O tables also were employed in a study of the construction industry authorized by the Economic Planning Board after publication of the Plan, in light of the 25 percent annual expansion rate being maintained by that industry. They were used again in the aforementioned integer programming study of the petrochemicals and iron and steel industries.¹¹

Some sectoral manpower requirements had been projected with the Second Plan's input-output model, but generally employment was neglected in the various planning studies, partly

¹⁰ Korean Development Association [9]. Professor Ronald McKinnon, of Stanford, and Dr. William Rapp, then of AID, developed the basic procedures for this exercise.

¹¹ The input-output tables are regarded as more reliable than the national accounts in Korea. The 1963 tables were used to revise the national accounts extensively. The tables also were tested for projection accuracy before their use in the Second Plan [2].

because the available estimates of employment levels were so poor and partly because the planners felt that labor would not be a constraint to growth during the Second Plan period. The latter assumption is being revised, particularly with respect to skilled labor, and employment studies will play a more significant role in the formulation of the Third Plan.

V. Concluding Remarks

Consistency, not optimality, has been the theme of the analytic work. This cautious approach means that planners were unprepared for some of the problems of rapid growth. Exports are a case in point.

Estimates of future export levels, viewed as projections by foreign planning advisers, were adopted as targets by the growth-conscious Korean officials. There is little doubt that Koreans have been willing to make some sacrifices in static allocative efficiency to achieve the dynamic stimulus provided by exports. To insure that the export targets, which have been more than doubled since the Second Plan was published, are met or exceeded, sectoral targets are further subdivided into industry targets, and in some cases into firm targets. A variety of incentives is offered to exporters, the principal ones being low interest rates on bank loans and exemption of duties on imports for reexport. A state trade promotion corporation has played an active role in securing foreign markets.

Given the Second Plan's concern with consistency—were reasonable growth rates indeed possible?—the possibility of significant inefficiencies in the export promotion measures was not seriously considered. However, the consequences of subsidized exports are beginning to appear; for example, the import content of export goods appears to be increasing, and some firms have raised prices on their domestic sales in order to offset the export losses.

Clearly, one of the major tasks of the Third Plan will be an assessment of the trade-off between the benefits of higher export growth and the costs in allocative inefficiencies.

At this point, a reevaluation of the planning process is under way as a prelude to the Third Plan work. The energy and command of the Second Plan effort has gradually been dissipated as the original Plan targets went through successive revisions. As the Plan expanded into reality through the annual ORB's, the focus became more selective. The exigencies of the budgeting process reasserted themselves.

With the Third Plan formulation, a reevalua-

tion and renewal of administrative momentum will occur. Due to the extensive experimentation of the Second Plan process, the approach to planning will be vastly different than it was five years earlier.

Perhaps the most important overall change is that quantitative tools are taken for granted, and they will be used in more refined ways, with more emphasis on methods of Plan implementation. There now is a stronger regional flavor to planning and more awareness of financial constraints. The push for basic industrialization is succeeding, and more attention will be devoted to technologically more sophisticated industries such as capital goods. Labor productivity and comparative advantage studies are being stressed.

There is a growing realization that growth is not sufficient as a social goal—that income distribution and social welfare issues must be confronted explicitly. But it must be acknowledged that the country's capability to resolve problems of a distributional nature is vastly enhanced because of the success of the growth strategy over several years. Growth per se or even higher living standards for all the people has not been the sole motivation for planning in Korea. For the planners in this non-Western, nonaffluent society, the vision of their world in 1981 is as follows: "For a typical family of five people, per family income will be 300,000 won per year, and thus will not only satisfy the basic necessities of life but will afford some margin for creativity and self-reflection in personal life" [12, p. 10].

REFERENCES

1. IRMA ADELMAN, ed., *Practical Approaches to Development Planning: Korea's Second Five-Year Plan* (Johns Hopkins Press, 1969).
2. IRMA ADELMAN, DAVID C. COLE, ROGER NORTON, AND LEE KEE JUNG, "The Korean Sectoral Model," Chap. 5 in Adelman [1].
3. BERNARD CAZES, "French Planning," Chap. 8, in Bert Hickman, ed., *Quantitative Planning of Economic Policy* (Brookings Institution, 1965).
4. JOHN G. GURLEY, HUGH T. PATRICK, AND E. S. SHAW, "The Financial Structure of Korea," USAID/Korea, July, 1965 (reprinted by the Bank of Korea).
5. KIM HAK SO, TARIK CARIM, AND R. D. NORTON, "Toward Development of Regional Planning Activities in Korea" (Seoul, Nov., 1967).
6. MAHN JE KIM AND R. D. NORTON, "A Stabilization Model for Korea," paper presented to the second Far Eastern meetings of the Econometric Society, Tokyo, June 29–July 1, 1967.
7. MAHN JE KIM AND DUCK WOO NAM, "A Statistical Model for Monetary Management: The Case of Korea, 1956–67," paper presented to the third

- Far Eastern meetings of the Econometric Society, Tokyo, June 27-June 29, 1968.
8. E. S. KIRSCHEN AND L. MORISSENS, "The Objectives and Instruments of Economic Policy," Chap. 5, in Hickman, *Quantitative Planning of Economic Policy*.
 9. Korean Development Association, *Effective Rates of Protection in Korea* (Seoul, 1967).
 10. HEE IL LEE, "Project Selection and Evaluation: Formulation of an Investment Program," Chap. 9, in Adelman [1].
 11. R. D. NORTON, "Formal Approaches to Regional Planning in Korea," Chap. 7, in Adelman [1].
 12. Republic of Korea, *Second Five-Year Economic Development Plan* (Seoul, July, 1966).
 13. Republic of Korea, *Overall Resources Budget: 1967* (Seoul, Jan., 1967).
 14. Republic of Korea, *Overall Resources Budget: 1968 (First Draft)* (Seoul, Mar., 1967).
 15. Republic of Korea, *Overall Resources Budget: 1968* (Seoul, Sept., 1967).
 16. Republic of Korea, *Overall Resources Budget: 1969* (Seoul, Sept., 1968).
 17. LARRY E. WESTPHAL, "Multisectoral Project Analysis Employing Mixed Integer Programming," Chap. 6, in Adelman [1].

QUANTIZED PREFERENCES AND PLANNING BY PRIORITIES

By THOMAS VIETORISZ
New School for Social Research

The objective of this paper is to explore the logic of project planning by priorities and to relate these to planners' preferences. Priorities, in the simplest case, are derived independently of resource-allocation considerations and override payoffs defined in terms of resources. The assumption of complete independence is later relaxed, and a mixed system making use of both priorities and material payoffs is sketched out.

The priority-based approach was suggested by problems of planning for the engineering industries in developing countries¹ and problems of defining an investment program for urban ghettos such as Harlem.² In both cases, in spite of the obvious relevance of resource allocation, a number of overriding pragmatic considerations could not be fitted into the format of conventional mathematical programming models. In the engineering industries, the role of a project as a potential focus for skill training and technological change, and in ghetto development, key payoffs in terms of community morale and political power, remained unquantifiable in terms of resources.

When such considerations dominate a decision problem, one might wish to turn to utility analysis. Yet no way exists for empirically ascertaining planners' (or consumers') utility functions. The method of revealed preferences is a mental experiment; the few authors who have tried to relate planners' preferences to quantitative models³ have

merely dealt with curvatures of indifference curves; while the literature on intertemporal consumption-investment decisions relies on a hypothetical discount rate as a proxy for time preference. Although in Eastern European discussions of mathematical planning models⁴ we find concern for what ought to be maximized, this concern, expressed in policy debates, has not been translated into new methods for including unquantifiable and extra-economic considerations in quantitative programming models.

Yet in a planning context, priorities suggest themselves as a well-tested means of expressing preferences that result from many imponderables. Priority orderings of major developmental alternatives were a prominent feature of planning in the U.S.S.R. since the first five-year plan, and production targets for selected industrial branches and key products, referred to as "leading links," have at times received overriding priority.⁵ The priority-based allocation of three key metals—steel, copper, and aluminum—was also hit upon after some false starts as the successful organizing principle of U. S. wartime planning.⁶ Priority orderings, as their past use in planning demonstrates, are operationally ascertainable—probably because, unlike indifference curves or surfaces, they are one-dimensional. In the sequel, planners' indifference curves defined in terms of resource flows will be constructed from two classes of simpler underlying information: the same planners' project priorities—which will be regarded as ascertainable and will not be further analyzed—and the resource inputs of alternative project variants.

Decision Modules and Project Variants

The operational use of priorities presupposes that the decision process be quantized; i.e., that it be undertaken in discrete units of finite (not infinitesimal) magnitude, capable of being numbered by a finite sequence of integers under a

¹ T. Vietorisz, *The Planning of Production and Exports in the Metalworking Industries*, UNIDO, Vienna, ID/WG.10/1, 1969; and *The Engineering Industries* (forthcoming UNIDO monograph).

² T. Vietorisz and B. Harrison, *The Economic Development of Harlem* (Praeger, in press).

³ J. Tinbergen in *Economic Policy: Principles and Design* (4th rev. printing, 1967) gives an aggregate model involving the tradeoff between employment and the balance-of-payments deficit, optimized by reference to hypothetical "indifference curves of a social welfare function." (Pp. 11, 101.) A. Theil, in *Economic Forecasts and Policy* (rev. ed., 1961) takes off from the same model to probe the analogy between planners' and consumers' preferences, then goes on to analyze in more detail the course of the model's indifference curves (pp. 374-79, 409). Chenery and Bruno, in "Development Alternatives in an Open Economy," *Econ. J.*, Mar., 1962, implement the trade-off curve empirically for Israel, and follow up by a very reserved qualitative discussion of the probable properties of the indifference curves of a hypothetical social welfare function (pp. 99-101).

⁴ For reference to this literature, see UNESCO International Bibliography on Economics (Annual).

⁵ See UN, Economic Commission for Europe, *Economic Planning in Europe*, Part 2 of *Economic Survey of Europe in 1962*, Chap. II, pp. 15-16.

⁶ See D. Novick, M. Anshen, and W. C. Truppner, *War Time Production Controls* (New York, 1949); and T. Schtovsky, E. Shaw, and L. Tarshis, *Mobilizing Resources for War* (New York, 1951).

given priority ordering.⁷ This creates no difficulties for planning decisions, since the most convenient unit of project planning is in any case the complex of interrelated activities. Treating linked⁸ activities as a group reduces the information needs of describing the structure of production and internalizes external economies and diseconomies,⁹ including those that result from nonconvexities (indivisibilities, economies of scale).¹⁰ A development plan is thus best constructed of activity groups (complexes) used as building blocks or modules. Each module is either included in a program or not; it cannot be used at half-scale or at other partial scales. This is the sense in which the decision-making process is said to be quantized. From this point on, the term "development project" will designate such activity complexes or modules.

In order to relate the quantized decision process to resource allocation models, it will be assumed that each project (complex, module) exists in several secondary variants that differ among themselves in regard to the intensity of capital, labor, or raw material use, the degree of mechanization or automation or other resource-bound characteristics. Ordinary resource allocation models choose among alternative projects and their alternative secondary variants at the same time, in such a way as to make the most efficient use of scarce resources. Just one major modification will be introduced into this scheme. A rank, designating the relative priority of the project, will be associated with each project, the rank being identical for all secondary variants. It is assumed that projects will be implemented in rank order, regardless of considerations of resource allocation. These considerations will be allowed to guide only the selection of secondary variants. If the range of such variants is broad, however, this still leaves much discretion for the efficient construction of development programs. Unless the choice of secondary variants of high-priority projects is efficient, an excessive amount of resources will be preempted and lower priority projects will either not be implemented or will be implemented only

with delay as additional resources become available in the course of time.

Graphical Analysis

Figure 1 represents the essential features of this problem. The coordinate axes measure the inputs of two material resources, f_1 and f_2 . Variants 1A and 1B of project 1 (of highest priority) are shown by vectors OP and OQ in Figure 1a. For expositional convenience the variants of each project are shown as the convex hull of two vectors, to be referred to as "extreme" variants. While it is postulated that the project must be executed as a whole, the extreme variants may be averaged to create new variants.

The joint resource requirements of two projects executed in rank order can be derived, as in Figure 1a, by fitting the tail of each vector representing an extreme variant of project 2 (2A or 2B) to the tip of a vector representing an extreme variant of project 1 (1A or 1B). The four possible combinations define the convex hull $RSTU$ of all feasible combinations of variants of the two projects. Yet only points along the line RST are efficient, since other points in $RSTU$ are dominated by some point in RST in terms of resource use. For adding the resource requirements of project 3 it is thus sufficient to start with RST , and so on for projects of progressively lower priority. The locus of efficient points representing the alternative resource requirements of the first n projects will be termed the n th indifference curve associated with the exogenously given priority ordering.

An objective function may be defined as $z = w_1f_1 + w_2f_2$, where w_1 , w_2 are factor prices and z is interpreted as total factor cost. The minimization of z for the first n projects can then be represented by the corner tangency between the n th indifference curve and a line having the slope $(-w_2/w_1)$. For $n=5$, the corner tangency occurs at X in Figure 1b. Point X represents the summation of project variants 1B, 2B, 3A, 4B, and 5B. The entire indifference curve can be traced out by parametrically varying $(-w_2/w_1)$.

By the commutative and associative laws of vector addition, the derivation of a given indifference curve is not dependent on adding up the project vectors in rank order; the order of addition is arbitrary. The essential properties of a priority ordering are therefore not contained in any single indifference curve, but only in the full sequence of these curves. This is seen in the converse of the above minimization problem where the task is to maximize the rank n of the indifference curve that can be reached subject to a prescribed budget constraint z . In passing from an

⁷ The logic of one-dimensional priority orderings might be capable of generalization to a sequence of discrete units whose size, in the limit, approaches zero.

⁸ Forward and backward linkages have been defined and extensively discussed by A. O. Hirschman in *The Strategy of Economic Development* (New Haven, 1958), Chaps. 6-8.

⁹ See T. Scitovsky, "Two Concepts of External Economies," *J.P.E.*, Apr., 1954.

¹⁰ See T. Vietoriz, "Decentralization and Project Evaluation in the Presence of External Economies and Diseconomies," *Industrialization and Productivity*, Bulletin 12, pp. 25-58, UN (New York, 1968).

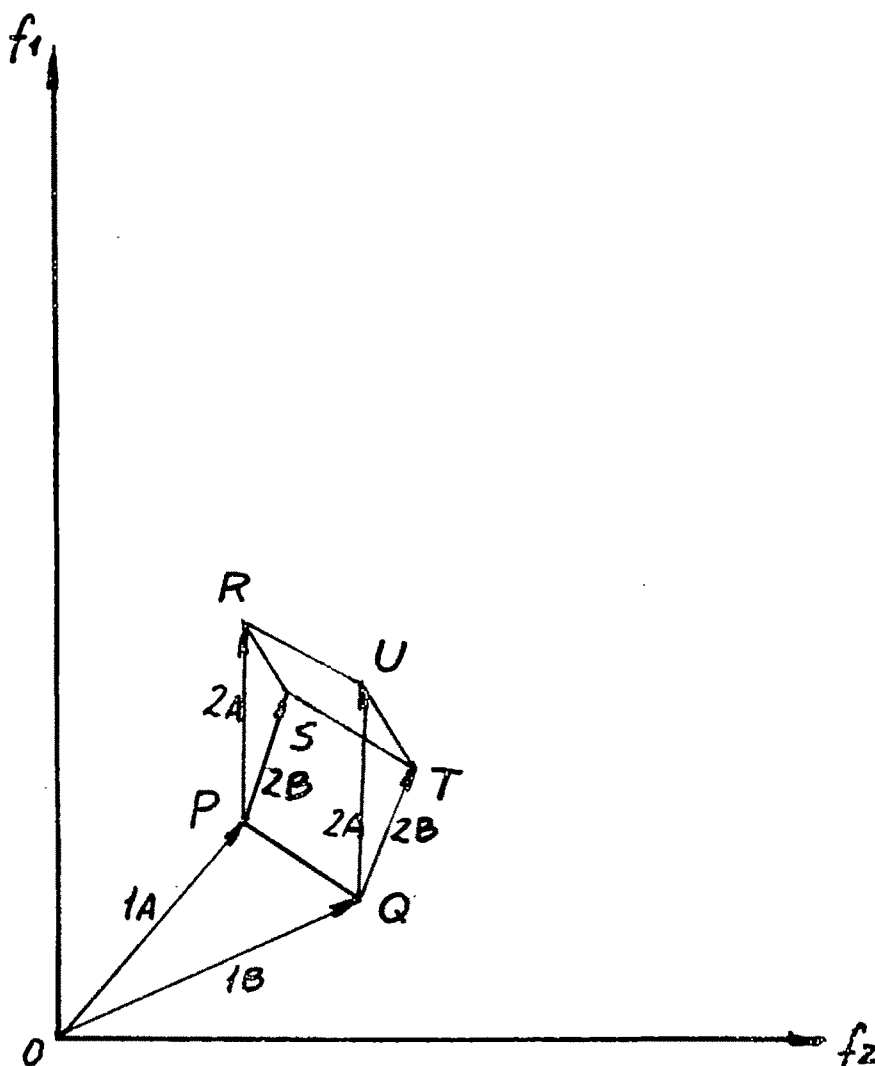


FIGURE 1a

indifference curve n to the next one, $n+1$, it is specifically the next project down the priority list that is to be added on.

As n increases only in integer increments, the maximization of n may leave part of the available resources unutilized, as when z is set slightly lower than the value corresponding to point X (see Figure 1). The maximal value of n then drops from 5 to 4.

The preference maps derived from priorities possess all the properties of ordinal utility functions except that they increase by discrete increments rather than continuously. Indifference curves are associated with ordered project sets whose utility increases as they include additional

projects further down the priority list. Resource inputs are thus valued not for themselves, but for what they contribute to the intermediate entities represented by the project sets. The derivation of these intermediate entities by linear transformations of resource inputs parallels Lancaster's derivation of "abstract qualities" which play an analogous intermediate role in his revision of consumer demand theory.¹¹ Where, however, Lancaster postulates an exogenous indifference map defined on "abstract qualities," the present approach starts with the simpler postulate of a one-dimensional priority ordering of the inter-

¹¹ Kelvin Lancaster, "A New Approach to Consumer Theory," *J.P.E.*, Apr., 1966, pp. 132-57.

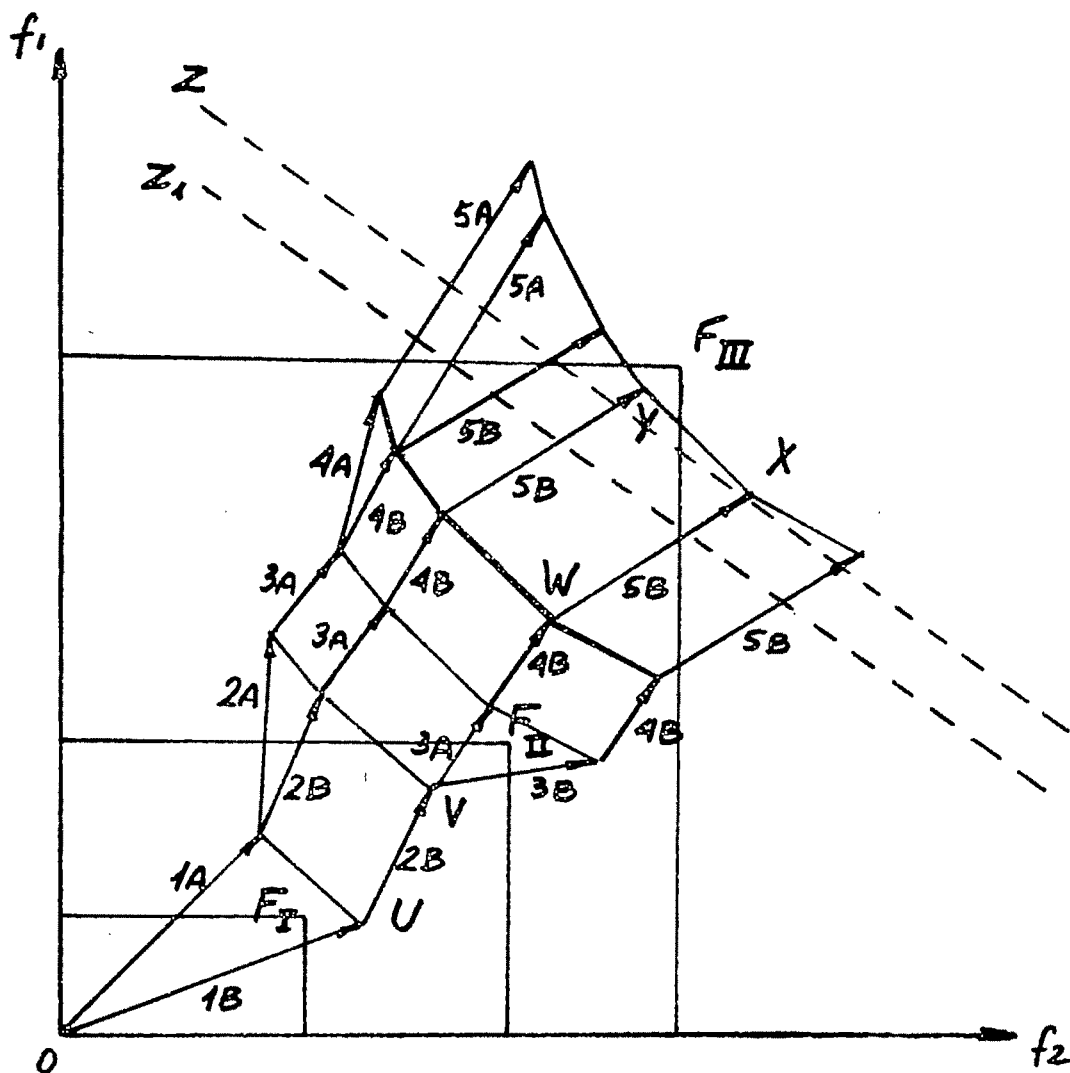


FIGURE 1b

mediate entities (the project sets) and proceeds to construct from these a preference ordering on the original resource input combinations.

A Dynamic Extension

So far the model has been entirely static; yet priorities, by their very nature, involve the passage of time. Assume that the exogenous supply of factors f_1 and f_2 expands over three time periods as indicated by points F_I , F_{II} , and F_{III} in Figure 1b. Note that a program constructed by optimizing over a long time horizon differs from one that evolves by a series of short-term optimizations. Thus the maximization of n , the number of projects implemented, over a three-year period leads

to point Y; while the maximization of the number of projects implemented in each time period, regardless of later consequences, leads to an evolution over the path O-U-V-W and attains only indifference level 4 over three periods rather than 5, as at Y. An inspection of Figure 1b will disclose that the implementation of project 5 in period III, under a long-horizon optimization, is made possible by a one-period postponement of the implementation of project 2.

Programming Formulation and Sensitivity Analysis

The formulation of the problem by means of a mathematical programming model involves the

maximization of the number of projects to be implemented, in priority sequence, subject to resource and/or budget constraints. While the sequencing constraints involve integer variables, the resulting integer programming problem can best be solved by restating it as a chain of linear programming problems in which the scale of the first n projects is set to unity, with n treated as a parameter that can assume values from 1 to N , the number of projects. This reformulation has the additional advantage that it yields shadow prices that can be readily interpreted.

In setting the first n project scales to unity, these scales are defined as the sums of the scales of the secondary variants of each project. The scale of the $(n+1)$ -th project is treated continuously and maximized, while the scales of lower-priority projects are set to zero. Trial and error readily establishes the correct value of n . If chosen too high, no feasible solution will exist; if too small, the scale of the $(n+1)$ -th project will exceed unity, contrary to assumption.

The Resource Cost of a Project Priority. When the scale of a high-priority project is forced to unity, the corresponding constraint has a shadow price that can be interpreted as a subsidy to the project, enabling it to break even at the resource prices of the optimal solution. The numerical value of the subsidy indicates by how much the scale of the $(n+1)$ -th project could be expanded if project i could be dropped in priority at least to rank $n+2$.

It may happen that a subsidy turns out to be negative; it can then be interpreted as a rent. This will occur when a project produces some scarce intermediate resource whose value exceeds the valuation of the input resources used up. The optimization will then attempt to raise the scale of this project above unity and will be restricted from so doing. This constraint produces a rent.

The Resource Cost of the Entire Priority Sequence. If the exogenous resource availabilities are expanded parametrically, e.g., following a time function of cumulative resource supply, then—using the above method—the cost of the first project can be expressed in terms of the second, the cost of the first and second in terms of the third, etc., by computing the respective project subsidies. Therefore, the implications of a given priority ordering can be made fully explicit: it can be shown what will be the consequences of assigning high priorities to certain projects, in terms of the implementation of projects of lower priority. This opens the door to an after-the-fact, iterative revision of initially assigned priorities by policy-makers, based upon the sensitivity of an

entire development program (consisting of a sequence of projects) to the priority ordering of projects within the sequence.

Alternative Formulation. Instead of maximizing the scale of the $(n+1)$ -th project, each project (and variant) can be given a payoff in terms of material resources, and the total payoff can then be maximized. As before, the scale of the first n projects is set to unity. The choice of secondary project variants will now depend on the exogenous payoff coefficients. Project subsidies are interpreted as the increase of overall payoff obtained when project i is dropped at least to rank $n+2$. Under this formulation, a priority can be replaced by an equivalent subsidy that will ensure that a project i is included in the set of the first n projects even if it appears unattractive in terms of material resource payoffs alone.

A Mixed System of Priorities and Resource Payoffs. The initial assumption that priorities override all resource-allocation considerations can now be relaxed. The computation of subsidies that are equivalent to the imposed priority reduces priorities and resource-based payoffs to a common denominator. This permits the definition of a mixed system in which the subsidies are subjected to constraints; for example, not to exceed 50 percent of the resource payoff coefficient. Since such constraints are formally part of the dual of the linear programming problem, they will have the effect of introducing a new primal activity which in turn introduces, at a specific cost level, a dummy replacement for the prescribed unit project scale. The effect is to create a feasible solution even though the project whose subsidy becomes excessive is dropped.

Conclusion

This paper indicates only summarily the potentialities of the proposed approach. The principal claim made on its behalf is that priority rankings are empirically ascertainable and that the resulting quantized preference maps can therefore be regarded as practically usable planning tools. They should be particularly useful in planning-programming-budgeting (PPB) tasks.

The approach, moreover, can be extended to other decision problems insofar as these are either naturally quantized or amenable to an approximation in these terms. Such an extension should open the way to empirical work in consumer demand theory, especially in the field of expenditures on durables and major services; and in the behavioral theory of the firm, especially in regard to expansion, diversification, and merger decisions.

A DYNAMIC FIVE-SECTOR MODEL FOR TURKEY, 1967-82*

By CHARLES R. BLITZER, HIKMET ÇETIN, and ALAN S. MANNE
Stanford University

I. Introduction

This paper reports upon a dynamic extension of the macromodel originally constructed as a consistency check upon the Second Five-Year Plan of Turkey. We address ourselves to plan formulation—a task that is comparatively free of ideological elements. We do not attempt a diagnosis of the existing structure of the Turkish economy, nor do we spell out the policy instruments (e.g., market incentives versus centralized controls) needed in order to steer the economy along the planned future trajectories.

Given the information base accumulated for the original consistency model, the dynamic extension turned out to be comparatively easy. No new data were collected, and only minor modifications were made in the technological norms. No more than a few man-months were needed in order to convert the original five-sector inter-industry system into a dynamic one—a model which could in turn be used to identify the areas where it would be fruitful to disaggregate further and to accumulate additional data. The ease of conversion is at least partially attributable to the formulation in terms of a “gradualist” consumption path. The gradualist path—together with certain additional hypotheses—permits us to adopt a short planning horizon for numerical computations (fifteen years), and yet to assert that the given plan would be not only feasible but also optimal if the planning horizon were extended over the infinite future. Aside from the objective function and terminal conditions, this model closely resembles those of Chakravarty and Lefebvre (1965), Chenery and MacEwan (1966) and Eckaus and Parikh (1968)—focusing on a labor-surplus economy, one that is currently a recipient of net foreign aid, and one that must plan for trade-balance-improving activities so as to increase its future political independence.

As in the original consistency model, the following five-sector interindustry classification is

employed here: (1) agriculture, (2) mining, (3) manufacturing, (4) construction, (5) services.

Our model covers a span of fifteen years, terminating in 1982 (the currently planned date for Turkey's entrance into full membership in the European Common Market). Sectoral balances are computed for the following representative periods—each a single year in length—and each spaced at intervals of 2.5 years:

<i>Representative Year</i>	<i>Time Period Index <i>i</i></i>
1967	0
1969-70 (average)	1
1972	2
1974-75 (average)	3
1977	4
1979-80 (average)	5
1982	6

II. The Maximand-Gradualist Consumption Paths

By restricting consumption paths to those of a gradualist pattern, we obtain a multisector formulation that is numerically computable and yet which retains something of the spirit of the Ramsey (1928) optimal savings model. The intertemporal choice is posed as one between consumption increases in the near future versus those in the distant future. Unlike the Ramsey formulation, it is required that all admissible consumption paths branch off from the known initial value C_0 . Moreover, asymptotically over time, it is required that consumption grow at the rate g . In general, the higher the value taken for the subjective policy parameter g , the more investment-oriented becomes the optimal development plan, and the lower the near-term rate of growth of consumption.

Let C_t denote aggregate consumption expenditures at date t . The quantity C_0 is a datum, the rate of actual consumption during the base year 1967. For subsequent years, the quantity C_t is evaluated simultaneously with the other unknowns of the programming model. Letting g denote the asymptotic growth rate (a subjective policy parameter), and letting D denote the initial consumption increment (a linear programming unknown), the formal definition of a gradualist path is as follows:

$$(1) \quad C_t = C_0 + D \left[\frac{(1+g)^t - 1}{g} \right], \quad (t=0, 1, 2, \dots, +\infty)$$

* Research supported in part by Ford Foundation International Studies grant to Stanford University and in part by U. S. Agency for International Development. The specific facts, methods of analysis, and conclusions are the sole responsibility of the individual authors. We are indebted to Kenneth Kauffman for helpful comments on earlier drafts. A more detailed version of this paper is available upon request to the authors.

It follows that:

$$D = C_1 - C_0,$$

and

$$g = \lim_{t \rightarrow +\infty} \left[\frac{C_{t+1} - C_t}{C_t} \right].$$

With $g > D/C_0 > 0$, note that $(C_{t+1} - C_t)/C_t$ (the percentage rate of growth of consumption) rises smoothly over time, asymptotically approaching g . The linear programming maximand is taken to be $D = C_1 - C_0$. Since the consumption increments in all other time periods are proportional to D , it follows that consumption is being maximized at all points of time—subject to the restriction imposed by equation (1) and to the fixed values assigned to C_0 and g . For our basic numerical calculations, we have set $g = 8$ percent, a quantity slightly higher than the 7 percent annual GNP growth rate target adopted officially for the Second Five-Year Plan.

The intertemporal trade-off is summarized in terms of the two parameters g and D . The higher the value of g , the lower that of D . Caution: It cannot be expected that two such parameters will be sufficient to characterize the development paths of all economies. All that is being proposed is an approximation that may prove useful during a labor-surplus development phase. For a mature economy, a more acceptable idealization would be that the asymptotic rate is governed by the growth of the labor force as measured in Harrod-neutral efficiency units. See, e.g., Solow, Tobin, Von Weizsäcker, and Yaari (1966).

III. Additional Basic Assumptions

For want of a detailed analysis of Turkey's agricultural sector, the output of that sector is assumed to grow at a fixed exogenous rate, 4.26 percent per annum. The flow of interindustry deliveries on current and capital account is assumed to be one directional—from the other sectors into agriculture. In this way, we allow for an acceleration in deliveries of nontraditional inputs (e.g., fertilizers and tractors) into agriculture—without supposing that agricultural inputs into industry (e.g., raw cotton) will grow as rapidly as manufacturing output itself.

For consistency with the assumption with respect to agriculture, the quantity C_t is defined as aggregate nonagricultural consumption expenditures during year t . Increments in nonagricultural consumption are to be delivered from the nonagricultural sectors in predetermined proportions: 1 percent from mining, 49 percent from manufacturing, and 50 percent from services. The asymptotic growth rate, together with the pre-

determined proportions for consumption increments, determines the parameters d_{it} , the increment in consumption demand for item i between period 0 and t , per unit of the maximand D .

As of the base year (1967), the bulk of Turkey's merchandise exports originated in the agriculture and mining sectors. Export earnings from these traditional items, together with invisibles, are projected exogenously throughout our planning horizon.

In 1967, virtually all of Turkey's merchandise imports consisted of manufactured products (sector 3). This is the sector in which there appears to be maximum scope for both import substitution and export promotion. Our investment planning model derives the requirements for trade-balance-improving activities by treating as an endogenous unknown Y_{it} , the imports less exports of manufactures during period t .

Re initial conditions: For the base year 1967, it is assumed that all quantities (output, inter-industry demands, consumption, etc.) are known except for the sectoral distribution of investment outlays. Subject only to the constraint that aggregate investment not exceed the known value of 17.58 billion TL during 1967 (period 0), we have supposed that the distribution by sector of destination is completely flexible. With this formulation, we err on the side of flexibility in the sectoral distribution of the initial increments in output.

Re terminal conditions: For the terminal year 1982 (period 6), it is supposed that the new capacity created will be in "turnpike" proportions, permitting subsequent investment growth to be maintained at the annual geometric rate $g (= 8.0$ percent for our basic case) in all sectors over the indefinite future. Although our formulation implies that asymptotically all sectoral capacities will grow at the identical rate g , it does not imply identical output growth in all sectors during the immediate post-terminal years.

IV. Algebraic Formulation

In this linear programming model, nonnegative values are to be assigned to 91 unknowns, defined as follows:

Definitions	Number of Unknowns
D = increase in nonagricultural consumption between period 0 and period 1 = maximand	1
X_{jt} = output increment in sector j between initial year and period t ($j = 1, \dots, 5$) ($t = 1, \dots, 6$)	30
Δ_{jt} = annual increment in capacity of sec-	

for j during 2.5 years centered around period t ($j=1, \dots, 5; t=0, 1, \dots, 6$)	35
Y_{it} = annual imports less exports of manufactures during period t ($t=1, \dots, 6$)	6
I_t = annual gross investment during period t ($t=0, 1, \dots, 6$)	7
S_t = annual domestic savings during period t ($t=1, \dots, 6$)	6
F_t = annual foreign loan inflow during period t ($t=1, \dots, 6$)	6
	—
	91

There are 91 equality and inequality constraints:

Identification of Constraint	Purpose	Number of Constraints
A	material balances	30
B	capacity constraints	30
C	terminal constraints on investment	5
D	definition of gross investment	7
E	relation between domestic savings, investment, and foreign loans	7
F	foreign exchange balance	6
G	upper bound on foreign loans	6
		—
		91

Since most of these 91 constraints are of a conventional nature, we provide details only with respect to (C), the terminal constraints on investment. These refer to the change in the material balance constraints between period 6 and 7. Let a_{ij} denote the current account input (if negative) or output (if positive) from sector i associated with a one unit gross output increase in sector j . Let b_{ij} denote the capital input requirement from sector i associated with a one unit increase in the gross annual output capacity of sector j . Then:

$$(C) \sum_{j=1}^5 [2.5a_{ij} - ((1+g)^{2.5} - 1)b_{ij}]\Delta_{jt} \geq (d_{i7} - d_{i6})D + \begin{bmatrix} \text{exogenous consumption} \\ \text{demand, period 7} \end{bmatrix} - \begin{bmatrix} \text{exogenous consumption} \\ \text{demand, period 6} \end{bmatrix} + \begin{bmatrix} \text{exogenous investment} \\ \text{demand, period 7} \end{bmatrix} - \begin{bmatrix} \text{exogenous investment} \\ \text{demand, period 6} \end{bmatrix} \quad (i = 1, \dots, 5)$$

For $t=6, 7, \dots, +\infty$, let:

$$(2) \quad \Delta_{j,t+1} = (1+g)^{2.5}\Delta_{jt},$$

and

$$(3) \quad X_{j,t+1} = 2.5\Delta_{jt} + X_{jt}.$$

If we suppose that the increments in consump-

tion and exogenous investment grow at the annual rate g , the terminal investment constraints (C), together with the primal solution (2) and (3), ensure that all material balance and capacity constraints will be satisfied over the infinite future following period 6.

Proof: For a period $t > 6$, multiply constraint (C) for item i by $(1+g)^{2.5(t-6)}$, and add to the corresponding material balance constraint (A) for item i , period 6. That the capacity constraints are also satisfied over an infinite horizon follows directly from (3) and the fact that constraints (B) are satisfied for period 6. This concludes the proof of primal feasibility for an infinite time horizon.

Note that this proof does not imply that the particular solution is optimal. For a statement of sufficient conditions to ensure optimality over an infinite planning horizon, see Hopkins (1969) and Manne (1970). These proofs hinge upon certain additional hypotheses concerning the optimal solution during the terminal periods of the finite horizon planning model: positive output and investment levels in all sectors, no slack capacities, and no excess production. Fortunately, these additional hypotheses are satisfied by the solutions recorded here, and so these solutions have the property of infinite horizon optimality.

V. Numerical Data

The numerical data were drawn from the latest available work sheets at the Turkish State Planning Organization and supplemented by our own informal estimates. Further work is needed to improve the reliability of these estimates.

VI. Numerical Results

Our five-sector model assumes that the growth of the Turkish economy will not be constrained by the growth of the labor force nor by labor productivity. Rather, we focus upon foreign exchange and capital accumulation constraints—as ex-

pressed in terms of two macroeconomic parameters: the asymptotic growth rate g , and the upper bound on foreign loan inflows. For the basic numerical results (Table 1), it is supposed that $g=8$ percent; that the limit on foreign loans will be 2.0 billion TL during periods 1, 2, and 3; that

TABLE 1

(Units: TL Billions, 1965 Prices)

 g = asymptotic growth rate of C = 8%

Year t	1967 0	1	1972 2	3	1977 4	5	1982 6
1. output of agriculture	38.80	42.51	47.37	52.75	58.69	65.24	72.35
2. output of mining	2.20	2.98	3.90	5.03	6.39	8.03	9.96
3. output of manufacturing	48.67	60.77	75.83	94.09	116.77	143.44	174.70
4. output of construction	8.62	10.74	12.98	15.76	18.76	22.37	26.14
5. output of services	53.53	60.73	69.77	80.25	92.98	108.20	126.50
Y_M , imports less exports of manu- factures	5.40	7.33	8.30	9.00	8.87	8.80	9.95
1. investment in agriculture	2.08	2.72	3.01	3.33	3.67	3.98	3.91
2. investment in mining57	.68	.83	1.00	1.21	1.42	1.63
3. investment in manufacturing	5.81	7.23	8.76	10.89	12.80	15.01	18.73
4. investment in construction76	.81	1.00	1.08	1.30	1.36	1.92
5. investment in services	4.90	6.15	7.12	8.66	10.35	12.44	14.50
Exogenous investment	3.46	5.05	7.00	8.75	10.80	13.30	15.90
I , Gross investment	17.58	22.63	27.74	33.70	40.12	47.51	56.59
F , Foreign loans	1.12	2.00	2.00	2.00	1.00	0.00	0.00
S , Domestic savings	16.46	20.63	25.74	31.70	39.12	47.51	56.59
TC , Total consumption	68.92	79.71	92.64	108.05	126.56	148.69	175.26
$GNP = TC + S$	85.38	100.34	118.37	139.76	165.68	196.19	231.84
Agricultural consumption	16.50	18.31	20.33	22.57	25.05	27.80	30.86
C , nonagricultural consumption	52.42	61.40	72.31	85.48	101.51	120.89	144.40
%/year growth of I		10.63	8.48	8.10	7.22	7.00	7.25
%/year growth of C		6.53	6.76	6.92	7.12	7.24	7.37
%/year growth of TC		5.99	6.20	6.35	6.53	6.66	6.80
%/year growth of GNP		6.67	6.83	6.87	7.04	6.99	6.91
Propensity to save (average for year 0; marginal thereafter)193	.279	.283	.279	.286	.275	.255
Incremental capital-output ratio		2.94	3.14	3.24	3.25	3.29	3.33

this limit will diminish to 1.0 billion during period 4; and that self-reliance will commence at period 5 and continue thereafter. Subject to the gradualist path restriction, consumption is to be maximized at all points of time

From Table 1, it can be seen that these macro-economic parameters imply growth rates slightly below those of the officially stated targets for the Second Five-Year Plan: 7 percent for GNP and 12 percent for manufacturing output. A closer approximation to the second plan targets is obtained if the value of g is raised to 10 percent. Along with this increase in g , there is a corresponding increase in requirements for fiscal austerity—as measured by the marginal propensity

to save. Figure 1 provides a visual comparison of the trade-off between consumption increases in the near versus distant future. With $g = 10$ percent rather than 8 percent, there would be a comparatively minor difference during the first three time periods. By period 4 (1977), however, the more austere policy would begin to yield additional consumption, and would provide an increasing advantage thereafter.

In another experiment, we left the asymptotic growth rate at 8 percent, and explored the implications of a substantial reduction in reliance upon foreign loans. The direct economic consequences of this move toward political independence were clear—a lowering of consumption targets during

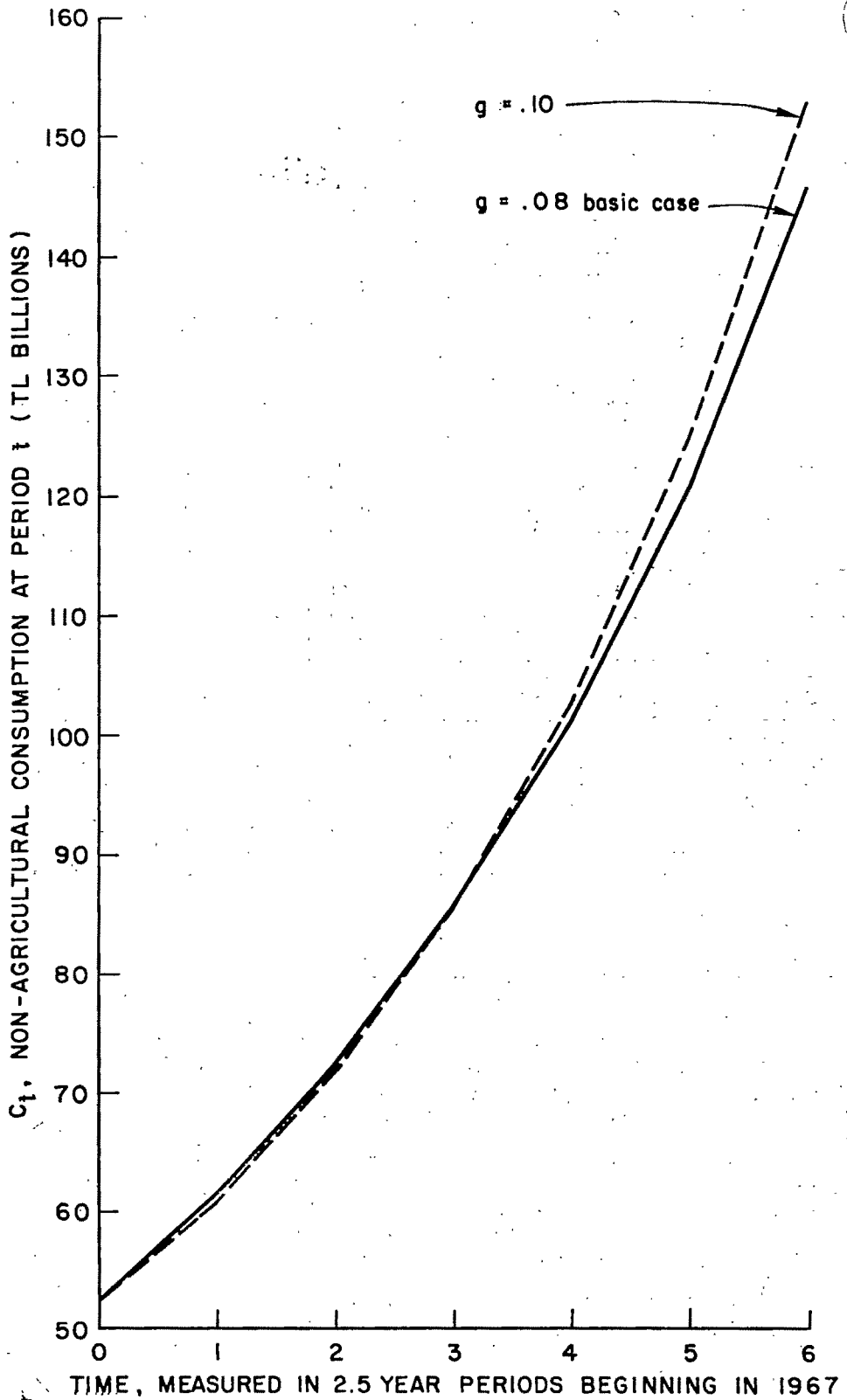


FIGURE 1

the Second Plan, an increase in the marginal savings ratio, and hence an increase in domestic austerity. Perhaps less obvious was the indirect effect—an increase in the Second Plan targets for manufacturing output; i.e., an increase in the requirements for trade-balance-improving activities.

Given a focus upon physical capital formation (ignoring labor constraints, education, nutrition, and human capital formation), it should come as no surprise that the marginal productivity of capital is of the same order of magnitude as the economy-wide output-capital ratio, 30 percent

per year. This is a point that has previously been emphasized by Harberger (1967, pp. 141-42) in his critique of project evaluation based upon a zero shadow price for labor. It is for this reason that we have performed alternative computations in which a portion of consumption demands is taken to be induced by factor payments. These computations lead to the suspicion that Table 1 overstates the productivity of physical capital and foreign loans. Moreover, the alternative shadow prices appear to provide more credible guidance for decentralized project evaluation.

AN EMPIRICAL TEST OF INTERREGIONAL INPUT-OUTPUT MODELS: ESTIMATION OF 1963 JAPANESE PRODUCTION*

By KAREN R. POLENSKE
Harvard University

As an economy develops, products sold in one region often are produced in another region of the country. This is one form of regional interdependence that can be analyzed within the framework of a spatially differentiated, general equilibrium trade model. For an analysis of regions within the American economy, one version of a large-scale multiregional input-output trade model will be implemented by the fall of 1970 at the Harvard Economic Research Project. The American economy will be separated into more than 30 regions and 60 industries. This will be the first time that a multiregional input-output model of this magnitude has been implemented for the United States. Many theoretical and empirical problems, however, must be solved before the model can be used for regional economic analyses on a routine basis.

The American model will incorporate a gravity trade model. The author implemented a small-scale version of the multiregional gravity trade model two years ago using interregional data from Japan [10]. Since the American data are still being assembled, the Japanese data are again used in this paper to compare the gravity trade model with two other spatially differentiated, general equilibrium trade models: a fixed column coefficient model and a fixed row coefficient model. Wassily Leontief, in collaboration with Alan Strout, tested a gravity trade model for individual commodity shipments [5]. Gravity trade models have also been incorporated within a general equilibrium framework to analyze transportation investment in underdeveloped countries [4] [11] to analyze regional production in Argentina [1] and Japan [10], and to study transportation requirements in what is called the "Northeast corridor" of the United States.¹ Chenery [2] and Moses [8] used the column coefficient trade model in their separate efforts to test empirically a multiregional input-output model. This author previously tested the row coefficient trade model

using fresh fruit and vegetable shipments [9], but this paper presents the first results of testing the row coefficient model within the overall input-output framework.

The three models tested are fixed trade coefficient models. A linear programming model, such as the one tested by Moses for the United States [7], would be an obvious alternative. Crosshauling of commodities,² however, cannot occur in linear programming models, and actual data on transportation costs are required. Since the aggregate nature of the interregional shipment data does produce crosshauls in the actual data and because consistent sets of transportation cost data are extremely difficult to obtain, a linear programming model was not included in the present study.

In comparison with the extensive data on transportation costs and other regional data required for a multiregional linear programming model, only a limited amount of actual regional data is needed to implement any of the fixed coefficient models. The required sets of regional data are: base-year technical coefficients, base-year trade coefficients, and a set of final demands for the given year. The technical and trade coefficients are assumed to remain fixed from year to year.

The Mathematical Formulation of the Multiregional Models

The steps necessary to obtain an operational model are only summarized in this paper since complete descriptions of the models have been presented in the articles referred to earlier in the paper. The notations for the models follow:

- $\hat{}$ indicates a block diagonal matrix
- Δ indicates the change between a base year and the given year
- n the number of regions
- m the number of commodities

The superscripts in this paper always refer to regions, while the subscripts always designate industries.

Matrix Notation

ΔX column vector ($mn \cdot 1$) giving the change in production. Each element describes the

² Crosshauling is the simultaneous shipment of the "same" good between two regions.

* This paper is financed with funds from Contract #7-35212 with the Economic Development Administration, United States Department of Commerce. Peter Solenberger did the programming and assisted with all phases of the research.

¹ The research on the Northeast corridor is being supported by the U.S. Dept. of Transportation, Transport Systems Planning Division.

- change in output of commodity i produced in region g .
- ΔY column vector ($mn \cdot 1$) giving the change in total final demand. Each element describes the change in the total amount of commodity i consumed by final users in region g regardless of the place where the good was produced.
- \hat{A} block diagonal matrix ($mn \cdot mn$) with n square matrices ($m \cdot m$) of input coefficients along the diagonal describing the structure of production in each region. If separate regional technical coefficients are not available, the matrix of national coefficients customarily is used for each region.
- S, T each is a square matrix ($nm \cdot nm$) filled with diagonal matrices ($m \cdot m$). The t_i^{gh} elements relate outflows from region g to the production in the region while the s_i^{gh} elements relate inflows into region h to the total consumption in the region.
- R square matrix ($nm \cdot nm$) filled with diagonal matrices ($m \cdot m$). Each r_i^{gh} element describes the fraction of total production of commodity i in region g that is exported to region h . The sum of each row of this matrix must equal 1, since the coefficients are proportions of total production.
- C square matrix ($nm \cdot nm$) filled with diag-

onal matrices ($m \cdot m$). Each element c_i^{gh} describes the fraction of total consumption of commodity i in region h that is imported from region g . The sum of each column of this matrix must equal 1, since the coefficients are proportions of total consumption.³

Element Notation

- x_i^{go} the total amount of commodity i produced in region g .
- x_i^{ah} the total amount of commodity i demanded by all final and intermediate consumers in region h .
- x_i^{**} the total amount of commodity i produced (consumed) in all regions.
- q_i^{gh} a trade parameter which is a function of the cost of transferring commodity i from region g to region h (where the transfer costs reflect various factors, including transportation costs, which determine interregional trade).

Some of the differences and similarities in the three models are shown by the summary in Table 1. In comparison with the row and column coeffi-

³ Because each industry within region h is assumed to consume the same fraction as imports: $c_i^{gh} = c_{ih}^{gh} = c_{ih}^{gh} = \dots = c_{im}^{gh}$. Isard suggested the use of the more detailed trade coefficients where the element c_{ij}^{gh} was the proportion of total consumption of commodity i purchased by industry j in region h which was produced in region g [3].

TABLE 1
SUMMARY OF THE MULTIREGIONAL INPUT-OUTPUT MODELS

	Row Coefficient Model	Column Coefficient Model	Gravity Model*
Trade coefficient equation	$x_i^{gh} = r_i^{gh} x_i^{go}$	$x_i^{gh} = c_i^{gh} x_i^{ah}$	$x_i^{gh} = \frac{x_i^{go} x_i^{ah}}{x_i^{**}} \cdot q_i^{gh}$
Equation system in matrix form	$R' \Delta X = \hat{A} \Delta X + \Delta Y$ $(R' - \hat{A}) \Delta X = \Delta Y$ $\Delta X = (R' - \hat{A})^{-1} \Delta Y$	$\Delta X = C(\hat{A} \Delta X + \Delta Y)$ $(I - CA) \Delta X = C \Delta Y$ $\Delta X = (I - CA)^{-1} C \Delta Y$	$T' \Delta X = S(\hat{A} \Delta X + \Delta Y)$ $(T' - S\hat{A}) \Delta X = S \Delta Y$ $\Delta X = (T' - S\hat{A})^{-1} S \Delta Y$

* The system of equations for the gravity model incorporates a simplified version of the basic gravity trade coefficient equation. The elements of S and T' in the gravity model are defined as:

$$s_i^{gh} = x_i^{go} \left[1 - \frac{x_i^{go} x_i^{gh}}{x_i^{**} x_i^{**}} \right] \quad \text{for } g \neq h \quad s_i^{gh} = 1 \quad \text{for } g = h$$

$$t_i^{gh} = x_i^{ah} \left[1 - \frac{x_i^{ah} x_i^{gh}}{x_i^{**} x_i^{**}} \right] \quad \text{for } g \neq h \quad t_i^{gh} = 1 \quad \text{for } g = h$$

where h can be assigned arbitrarily.

TABLE 2
ABSOLUTE ERRORS FOR ROW COEFFICIENT, COLUMN COEFFICIENT AND GRAVITY TRADE MODELS
JAPANESE 1963 OUTPUTS BY INDUSTRY AND REGION
(Billions of Yen)

Region	Industry										Total All Industries
	Agriculture Forestry Fisheries	Mining	Textiles	Chemicals	Metals	Machinery	Other Manufac- turing	Building and Con- struction	Transpor- tation	All Other	
1. Hokkaido.....	4.0 101.8 101.7 298.4	-315.3 24.4 38.0 106.0	-13.3 7.9 8.0 12.9	-89.0 10.7 11.9 61.8	-1020.2 -12.7 16.6 303.0	-43.7 26.0 25.6 46.3	-90.1 2.9 -3.8 492.2	-3.1 7.4 7.6 204.4	-39.7 0.1 2.1 126.7	-142.1 52.4 56.7 657.3	-1752.4 221.0 264.4 2309.0
2. Tohoku.....	186.5 88.9 57.0 528.4	612.0 13.5 4.3 56.1	-42.6 4.1 3.1 52.3	598.4 13.9 -1.3 104.4	1217.5 56.1 25.5 193.5	140.9 -5.6 -9.8 114.8	40.0 -14.7 -40.4 563.8	36.0 0.5 -0.7 280.8	99.3 16.3 14.9 118.8	468.1 44.3 27.0 750.8	3356.1 217.3 79.7 2763.6
3. Kanto.....	413.3 161.5 199.4 908.7	96.2 1.2 5.1 66.9	491.1 66.5 96.0 616.1	406.9 -195.3 -133.7 1176.1	1528.9 358.3 445.0 1758.6	539.9 -43.7 25.6 3621.2	552.0 76.8 170.8 3237.8	27.9 3.8 6.5 1889.8	90.2 11.7 21.0 961.2	548.1 -38.4 33.7 5361.7	4694.4 402.5 869.5 19598.1
4. Tokai.....	110.9 36.5 34.2 228.0	13.0 -0.1 -3.2 21.2	639.3 89.9 105.1 841.6	-119.9 -78.9 -85.0 375.5	-903.9 -70.6 -77.2 379.8	-513.6 -108.9 -133.5 888.3	-170.6 -70.0 -83.2 870.1	-8.3 -4.0 -4.3 519.5	-12.6 12.9 11.8 156.7	-126.1 -18.7 -25.9 1212.8	-1091.8 -211.8 -261.3 5493.4
5. Hokuriku.....	17.5 9.7 4.3 74.0	36.3 1.7 1.1 3.5	-68.1 21.1 16.3 134.6	575.5 8.5 8.9 60.6	-148.6 11.8 -2.3 86.3	56.6 -6.1 -11.8 105.4	25.3 2.1 -3.2 143.3	8.0 -0.0 -0.3 108.6	26.0 -0.9 -2.6 44.3	141.9 -4.1 -9.1 258.9	670.5 43.8 1.3 1019.6

TABLE 2 (Continued next page)

Region	Industry										Total All Industries
	Agriculture Forestry Fisheries	Mining	Textiles	Chemicals	Metals	Machinery	Other Manufac- turing	Building and Con- struction	Transpor- tation	All Other	
6. Kinki.....	-379.1 69.4 62.5 283.4	-199.7 1.0 -6.4 26.6	-1012.0 101.8 63.3 827.3	-2172.7 -32.3 -51.6 745.8	1754.4 76.5 100.1 2010.5	92.8 189.9 187.9 1506.5	-470.9 -109.4 -113.9 1832.4	-29.9 -1.3 -2.0 852.0	-65.2 27.7 24.9 469.4	-697.7 61.2 36.8 2752.1	-3180.1 384.7 301.6 11305.8
7. Chugoku.....	290.2 80.8 78.8 281.1	269.1 6.0 -6.0 28.0	401.9 -19.6 -16.4 230.2	453.4 -54.8 -60.7 439.2	-376.5 -41.0 -26.1 375.6	-8.7 -46.0 -61.6 430.6	130.8 -7.6 -15.9 592.6	18.0 2.4 1.8 303.8	32.8 3.0 0.9 127.2	230.1 24.7 16.6 729.6	1441.1 -52.0 -88.6 3538.0
8. Shikoku.....	71.3 43.8 33.6 185.8	-15.4 -0.3 -4.6 14.3	39.2 26.6 24.0 68.8	232.3 -49.2 -71.5 222.8	-174.6 9.0 -2.7 74.2	16.8 3.0 1.9 72.9	29.4 -8.8 -15.4 300.3	1.8 -0.2 -0.8 136.4	17.8 9.1 7.7 58.1	32.0 0.1 -8.3 379.5	250.6 33.0 -36.0 1513.1
9. Kyushu.....	-33.5 107.8 120.6 567.0	-378.0 58.4 69.4 118.5	-45.2 13.9 15.7 49.5	-233.6 43.0 45.9 255.9	-1352.8 400.2 332.5 828.9	-213.2 26.7 10.8 210.7	-85.5 75.0 57.0 798.7	-17.4 1.6 1.4 435.6	-51.8 19.3 17.2 182.7	-273.7 81.8 73.6 1210.1	-2684.8 827.7 744.0 4657.6
Total, all regions.....	681.3 700.1 692.0 3354.9	118.2 105.9 97.8 441.1	390.2 312.2 315.0 2833.3	-348.8 -334.4 -337.2 3442.1	524.1 787.6 811.4 6010.5	67.7 35.4 35.3 6996.6	-39.5 -53.6 -48.0 8831.2	33.0 10.2 9.2 4730.9	96.9 99.2 98.0 2244.9	180.5 203.5 201.0 13312.8	1703.4 1866.2 1874.6 52198.3

cient models, the interregional trade equation for the gravity model is nonlinear, although it is linearized in the actual implementation of the model. The gravity model also relies explicitly upon a reasonable assumption that changes both in production and consumption affect the interregional shipments of commodities. The similarities and differences among the three equation systems will be discussed in detail in a future paper.

*Comparison of the Japanese 1963
Regional Production Estimates*

The Japanese Ministry of International Trade and Industry (MITI) assembled a nine-region, interregional input-output table for 1960 and a set of nine intraregional input-output tables for 1963.⁴ For this paper, data from the 1960 interregional table were used to calculate the base-year regional technical coefficients and the base-year trade coefficients for ten industries in each of nine regions in Japan. The final demands were calculated from the 1963 intraregional tables. In 1960 over one-half of Japanese goods and services were produced in the two regions of Kanto and Kinki. Japanese industries, of course, are experiencing a rapid rate of growth. Over the four-year period from 1960 to 1963, the most substantial output increases occurred in three industries: chemical production increased by 59 percent, machinery production by 51 percent, and construction by 49 percent.

The three multiregional input-output models, described in the previous pages, were implemented using the sets of base-year data for 1960 and the 1963 regional final demands to estimate 1963 regional outputs and interregional trade flows. The accuracy of each model is determined by comparing the actual and the estimated regional outputs. An investigation can then be made to discover the economic conditions that influence the accuracy of each model.

Several measures are used to show the differences in the predictive power of each model. Table 2 gives the discrepancy between the estimated 1963 regional outputs, \hat{x}_i^{go} , and the actual 1963 regional outputs, x_i^{go} . These errors will be referred to as the absolute errors. If they are positive, the output is overestimated, and if they are negative, the output is underestimated. In

each set of four figures, the first three lines give the absolute error for the row coefficient, the column coefficient, and the gravity trade model estimates, respectively. The fourth line provides the actual output figures for comparison with the absolute errors. Thus, although the discrepancy between the actual and the estimated output is very large for the metal industry in Kinki (over 100 billion yen for the gravity trade model) the discrepancy represents less than 5 percent of total metal output in Kinki. The last column in the table shows the aggregate error (all industries) for each region. The last row gives the aggregate error (all regions) for each industry. In almost every set of four figures, the absolute error for the row coefficient model is noticeably larger than the errors for the other two models. The row coefficient model seems to do better when producers have strong marketing ties to major purchasers. When the three models were previously tested by the author [9] using one commodity, fresh fruits and vegetables, the errors of estimation for the row coefficient and gravity trade models were low while the errors of estimation for the column coefficient model were significantly higher.

The relative importance of the absolute errors in relation to total output for an industry in a region is shown by the percentage errors of estimation contained in Table 3. These are calculated as:

$$e_i^{go} = \frac{\hat{x}_i^{go} - x_i^{go}}{x_i^{go}} \quad i = 1, \dots, n \\ g = 1, \dots, n$$

The data for the three models are organized as for Table 2. In each of the nine regions, the aggregate errors (all industries) of estimation for the row coefficient model are much larger than those calculated for the other two models. The last row gives the aggregate percentage errors of estimation for each industry. The row coefficient model does provide slightly lower errors of estimation than the other two models for five of the ten industries: agriculture, metal, other manufacturing, transportation, and the industrial category entitled "all others." The small aggregate errors shown for the row coefficient model are illusory since they result after balancing large negative errors against large positive errors.

The 90 errors for each of the three models from Table 2 and Table 3 are tabulated in Table 4 to show the distribution of the errors from large negative figures to large positive figures. For both the column coefficient and the gravity trade models, thirty-three out of ninety production estimates are within 10 billion yen or less of the actual output. In terms of the percentage errors

⁴ Before making the calculations described in this paper, a few adjustments were made to balance the 1963 tables. All of the Japanese figures cited are obtained from the published set of tables. A more complete description of the regional economies in Japan is contained in a previous paper by the author [10].

TABLE 3
PERCENTAGE ERRORS OF ESTIMATION FOR ROW COEFFICIENT, COLUMN COEFFICIENT AND GRAVITY TRADE MODELS
FOR 1963 JAPANESE OUTPUTS BY INDUSTRY AND REGION
(Percent)

Region	Industry										Total All Industries
	Agriculture Forestry Fishes	Mining	Textiles	Chemicals	Metals	Machinery	Other Manufac- turing	Building and Con- struction	Transpor- tation	All Other	
1. Hokkaido.....	1.4 34.1 34.1	-297.5 23.0 35.9	-102.9 61.1 61.8	-144.1 17.4 19.2	-336.7 -4.2 5.5	-94.3 56.2 55.3	-18.3 0.6 -0.8	-1.5 3.6 3.7	-31.3 0.1 1.7	-21.6 8.0 8.6	-75.9 9.6 11.4
2. Tohoku.....	35.3 16.8 10.8	1090.1 24.0 7.7	-81.5 7.9 5.9	573.2 13.3 -1.3	629.1 29.0 13.2	122.8 -4.9 -8.5	7.1 -2.6 -7.2	12.8 0.2 -0.2	83.7 13.7 12.6	62.3 5.9 3.6	121.4 7.9 2.9
3. Kanto.....	45.5 17.8 21.9	143.8 1.9 7.7	79.7 10.8 15.6	34.6 -16.6 -11.4	86.9 20.4 25.3	14.9 -1.2 0.7	17.0 2.4 5.3	1.5 0.2 0.3	9.4 1.2 2.2	10.2 -0.7 0.6	24.0 2.1 4.4
4. Tokai.....	48.7 16.0 15.0	61.5 -0.3 -15.0	76.0 10.7 12.5	-31.9 -21.0 -22.6	-238.0 -18.6 -20.3	-57.8 -12.3 -15.0	-19.6 -8.0 -9.6	-1.6 -0.8 -0.8	-8.0 8.3 7.5	-10.4 -1.5 -2.1	-19.9 -3.9 -4.8
5. Hokuriku.....	23.7 13.1 5.8	1032.0 48.7 30.9	-50.6 15.7 12.1	949.2 14.0 14.7	-172.2 13.6 -2.6	53.7 -5.8 -11.2	17.7 1.5 -2.2	7.3 -0.0 -0.3	58.7 -2.1 -5.8	54.8 -1.6 -3.5	65.8 4.3 0.1
6. Kinki.....	-133.8 24.5 22.1	-751.0 3.9 -24.0	-122.3 12.3 7.7	-291.3 -4.3 -6.9	87.3 3.8 5.0	6.2 12.6 12.5	-25.7 -6.0 -6.2	-3.5 -0.1 -0.2	-13.9 5.9 5.3	-25.4 2.2 1.3	-28.1 3.4 2.7
7. Chugoku.....	103.2 28.7 28.0	961.2 21.5 -21.6	174.6 -8.5 -7.1	103.2 -12.5 -13.8	-100.2 -10.9 -7.0	-2.0 -10.7 -14.3	22.1 -1.3 -2.7	5.9 0.8 0.6	25.8 2.3 0.7	31.5 3.4 2.3	40.7 -1.5 -2.5
8. Shikoku.....	38.4 23.6 18.1	-107.9 -2.0 -32.4	57.0 38.7 35.0	104.3 -22.1 -32.1	-235.3 12.1 -3.7	23.0 4.1 2.6	9.8 -2.9 -5.1	1.3 -0.2 -0.6	30.7 15.6 13.3	8.4 0.0 -2.2	16.6 2.2 -2.4
9. Kyushu.....	-5.9 19.0 21.3	-319.1 49.3 58.6	-91.3 28.0 31.6	-91.3 16.8 17.9	-163.2 48.3 40.1	-101.2 12.7 5.1	-10.7 9.4 7.1	-4.0 0.4 0.3	-28.3 10.6 9.4	-22.6 6.8 6.1	-57.6 17.8 16.0
Total, all regions.....	20.3 20.9 20.6	26.8 24.0 22.2	13.8 11.0 11.1	-10.1 -9.7 -9.8	8.7 13.1 13.5	1.0 0.5 0.5	-0.4 -0.6 -0.5	0.7 0.2 0.2	4.3 4.4 4.4	1.4 1.5 1.5	3.3 3.6 3.6

TABLE 4
DISTRIBUTION OF ABSOLUTE AND PERCENTAGE ERRORS OF ESTIMATION

Model \ Range	Negative Errors						Positive Errors					
	50 and over	30-50	20-30	10-20	5-10	0-5	0-5	5-10	10-20	20-30	30-50	50 and over
							absolute (billions of Yen)					
Row coefficient.....	28	5	1	4	2	1	2	1	5	4	6	31
Column coefficient...	7	6	0	4	4	8	14	7	9	7	4	20
Gravity.....	10	1	2	4	5	13	9	6	9	7	6	18
							percentage					
Row coefficient.....	22	2	5	5	2	5	3	7	5	4	7	23
Column coefficient...	0	0	2	6	4	17	18	7	20	9	5	2
Gravity.....	0	2	4	6	9	14	14	14	13	5	6	3

of estimation shown in the bottom half of the table, nearly one-third of the production estimates for the two models are within 5 percent of the actual output value and over one-half are within 10 percent. The row coefficient model does not perform nearly so well. Substantially fewer estimates are within the plus or minus 10 percent error range, and some relative errors of estimation exceed 100 percent.

Conclusion

This paper is too short for a detailed presentation of all the test results. Substitution of 1963 technology for the 1960 base-year technology decreases the errors of estimation for most industries, but substantial errors still occur for some industries as a result of shifts in interregional shipments. Separate calculations were also made using the four versions of the gravity trade model formulated by Leontief and Strout [5]. Although the point estimate version requires more actual data than the other three, it also, usually, produces the lowest errors of estimation. Therefore, only the point estimate gravity trade model results are presented in this paper.

All of the tests made so far indicate that the row coefficient model estimates Japanese regional production less accurately than the column coefficient or gravity trade models. The method of handling foreign trade in the model may contribute to the instability of the row coefficients for Japan. For any of the three models, a separate estimation of interregional shipments of foreign exports and imports may be desirable because of the importance of foreign trade in Japan. Finally, the tests reveal no discernible difference in the overall predictive ability of the column coefficient and the point estimate gravity trade models. Both models give reasonable estimates of Japanese regional production.

REFERENCES

- MARIO S. BRODESOHN, *A Multiregional Input-Output Analysis of the Argentine Economy* (Instituto Torcuato di Tella, Centro de Investigaciones Económicas, Buenos Aires, Oct., 1965).
- HOLLIS CHENERY, "Regional Analysis," *The Structure and Growth of the Italian Economy*, ed., H. Chenery and P. Clark (Rome: United States Mutual Security Agency, 1953), pp. 97-129.
- WALTER ISARD, *Methods of Regional Analysis* (M.I.T. Press, 1963).
- DAVID KRESGE, "A Simulation Model for Development Planning," Harvard Transportation and Economic Development Seminar, Discussion Paper No. 32, Nov., 1965 (unpublished).
- WASSILY LEONTIEF and ALAN STROUT, "Multiregional Input-Output Analysis," *Input-Output Economics*, ed., Wassily Leontief (Oxford Univ. Press, 1966), pp. 223-57.
- MITI (Ministry of International Trade and Industry). "Inter-regional Input-Output Table for Japan," *Trade and Industry of Japan*, No. 108 (undated).
- LEON N. MOSES, "A General Equilibrium Model of Production, Interregional Trade, and Location of Industry," *Rev. of Econ. and Statis.*, Nov., 1960, pp. 373-97.
- , "The Stability of Interregional Trading Patterns and Input-Output Analysis," *A.E.R.*, Dec., 1955, pp. 803-32.
- KAREN R. POLENSKE, "A Case Study of Transportation Models Used in Multiregional Analysis" (unpublished Ph.D. dissertation, Harvard Univ., May, 1966).
- , "Empirical Implementation of a Multiregional Input-Output Gravity Trade Model," *Contributions to Input-Output Analysis*, ed., A. P. Carter and A. Bródy (Amsterdam: North-Holland Pub. Co.) (forthcoming in 1970).
- PAUL O. ROBERTS, "The Role of Transport in Developing Countries: A Development Planning Model" (a paper presented Dec., 1967, at the Pan American Highway Congress, Montevideo, Uruguay) (unpublished).

INDUSTRIAL ORGANIZATION: RETROSPECT AND PROSPECT

INDUSTRIAL ORGANIZATION: PAST HISTORY AND FUTURE PROBLEMS

By E. T. GRETHER

University of California, Berkeley

I. Retrospect

The label, "Industrial Organization," and the initial form and impetus came out of Harvard. At Harvard and elsewhere a variety of descriptive institutional courses had appeared in such fields as utilities, trusts, corporations, financial organization, agriculture, and marketing. Typically, these courses were not integrated with economic theory, perhaps, because of the high interest in and need for collections of factual materials. The Great Depression and the NRA experience and the publication of Berle and Means, *The Modern Corporation and Private Property*, in 1932, led to a demand for a more basic theoretical-empirical approach. The hearings and studies and final report of the Temporary National Economic Committee on the Concentration of Economic Power beginning in 1938 provided additional emphasis and welcome empirical materials.

The first phase at Harvard definitely reflected "the Chamberlinian Revolution"; literally so, because the initial exploratory course work succeeding the traditional course on corporations, which had been given by W. Z. Ripley, was a joint effort of Professors Edward H. Chamberlin and Edward S. Mason. The empirical-institutional and public policy pulls were stronger for Mason than for Chamberlin, for he continued to take greater strides into the real world with the realization that this was to some extent at the expense of theoretical elegance. In his general introduction to the collection of essays written over the period 1936-56 and published in 1957 under the title, *Economic Concentration and the Monopoly Problem*, Mason characterized the study of business organizations as "eclectic methodologically" and as a "muddy, but not uninteresting, field."¹

Gradually, at Harvard, a corps of coworkers arose, including scholars in law and public administration, in addition to economists. Among the very first was Donald H. Wallace, whose study of market control in the aluminum industry became the model for other investigations.² By

1938, Professor Mason was able to present the framework of analysis as then envisaged before this Association at its meetings in Detroit, Michigan.³

The primary interest as outlined in the 1938 position paper was in price policy: "the deliberative action of buyers and sellers to influence price" and especially the policies of large-scale industrial enterprises.⁴ Although there was a heavy empirical-institutional stress, it was held that "a frame of reference of important greater generality than that of the institutionalist was required for useful work."⁵ The conceptual orientation was from the theories of monopolistic competition and oligopoly with the internal organization and decision making of enterprises inherently and explicitly an important part of the framework but in the setting of the market structure of each enterprise. Consequently, markets and market structures, so it was stated, "must be defined with reference to the position of a single seller or buyer."⁶ Finally, and in logical sequence, the primary contribution to "greater generality" would be through the classification or grouping of firms in terms of similar market structures and market conditions with the hope and expectation that "a careful study of the empirically determinable differences in market structure may go far in explaining observable differences in policy and practice."⁷

Consequently, a basic aspect of the approach was and continues to be the conceptualization and definition of market structures. So far as I can discover, the verbiage, "market structure," arose out of the discussions in the original Harvard group. Very early, however, a difference of viewpoint arose which still continues. In his 1938 position paper, Mason stated explicitly:

... The market, and market structure, must be defined with reference to the position of a single seller or buyer. The structure of a seller's market, then, includes all those considerations which he takes into account in

¹ E. S. Mason, "Price and Production Policies of Large-Scale Enterprises," *A.E.R.*, Mar., 1939, sup. Reprinted at pp. 55-72 in Mason, *op. cit.* Page references here are to the reprint, not the original.

² *Ibid.*, p. 55.

³ *Ibid.*, p. 55.

⁴ *Ibid.*, p. 65.

⁵ *Ibid.*, p. 66.

¹ Edward S. Mason, *Economic Concentration and the Monopoly Problem* (Harvard Univ. Press, 1957), pp. 4 and 8.

² D. H. Wallace, *Market Control in the Aluminum Industry* (Harvard Univ. Press, 1937).

determining his business policies and practices. His market includes all buyers and sellers, of whatever product whose action he considers to influence his volume of sales.⁸

The grouping together of firms for purposes of analysis, however, would be in terms of other factors than market structure alone, unless the concept of structure were extended to include these factors or conditions. Mason listed five types of such conditions: (1) the economic characteristics of the product; (2) cost and production characteristics of the firm's operations; (3) numbers and relative sizes of buyers and sellers and relative ease of entry of new firms; (4) demand conditions; and (5) differences in distribution channels. Admittedly, too, this was an incomplete listing; each situation must be investigated in the search for the relevant conditions.⁹ Although the subjective and enterprise goal factors affecting decision-makers were explicitly involved, the stress definitely was upon the objective and external elements of the environment of decision making. Thus, it was stated that in the tire industry, "while the personality of Firestone, plus the fact that his firm is admittedly a low-cost producer, has no doubt been an important factor, it seems probable that if Firestone, like God in another context, had not existed, the structure of the tire market would have created him."¹⁰

I shall not recall the variations of emphasis and usage over the years. Fortunately, Joe S. Bain surveyed the literature down to 1947. Hence, I may leave to your rereading his summaries of empirical materials and findings. He noted a surprising amount of research evidence. He concluded, however, "in spite of the aggregate, resultant contribution, our empirical knowledge of price making, its origins, and its results is still fragmentary."¹¹

Bain also discussed at some length the issue of objective determinism (referring to the Mason-Nourse polemic) through market structure analysis and the use of market structure characteristics as bases of more refined market classifications. He presented, for example, an abbreviated classification based upon a model introduced by E. S. Mason a decade before.¹² But he concluded that the industry studies and other empirical researches have made "little definite progress as

yet in establishing an objective classification of markets, each subcategory of which would contain industries with a uniform and distinctive type of competitive behavior."¹³ In this connection, he reviewed and appraised briefly a number of classifications and criticized them as either nonobjective, or too general, or lacking empirical support, or in some cases, even the possibility of empirical testing.¹⁴

Exceedingly interesting, in terms of the issue of objective determinism through the market structure-market classification framework of analysis, was the tentative suggestion that at any given time (or under a static or quasi-static market structure) executive market discretion might be confined to a relatively narrow range; but over time the executive has enough discretion as to allow "significantly different policies designed to change market structure." Consequently, "the dynamic course of market structure (and hence behavior) over substantial time intervals may not be at all determinate."¹⁵ Bain also raised the issue of the scope and focus of investigations and concluded that both the firm and the industry are appropriate foci, but insisted that the investigations must be intensive case studies. Finally, Bain lamented the lack of "an adequate and dependable set of norms of satisfactory price-output results for individual industries"¹⁶ and stressed the need for the development of "a basic theory and of norms appropriate to a quasi-monopolistic economy in dynamic process."¹⁷

It is important to note, too, that Bain, who was one of the earliest of Mason's Ph.D.'s in this field, in his own work, clearly and definitely decided to stress the traditional industry, partial equilibrium approach¹⁸; and this pattern has become the more characteristic and productive one in industrial organization to this point. The industry approach has numerous advantages, including the recognizable line of descent from neoclassical partial equilibrium theory and the ability to use Census data developed in terms of natural categories. A substantial number of useful industry studies have been made in this context usually by the use of Census and other industry data.¹⁹ An

⁸ *Ibid.*, p. 65.

⁹ *Ibid.*, pp. 65, 66.

¹⁰ *Ibid.*, p. 68.

¹¹ J. S. Bain, "Price and Production Policies," in *A Survey of Contemporary Economics*, H. S. Ellis, ed. (Richard D. Irwin, Inc., 1948), p. 151.

¹² *Ibid.*, p. 161.

¹³ *Ibid.*, p. 158.

¹⁴ *Ibid.*, pp. 158, 159.

¹⁵ *Ibid.*, p. 157.

¹⁶ *Ibid.*, p. 169.

¹⁷ *Ibid.*, p. 170.

¹⁸ J. S. Bain, *The Economics of the Pacific Coast Petroleum Industry*, Part I: Market Structure, p. 11, fn. 3.

¹⁹ See Bain, above, for a listing of studies down to 1947. Among the published book length studies since then are: Melvin G. de Chazeau and Alfred E. Kahn, *Integration and Competition in the Petroleum Industry*

even more numerous body of cross-industry studies on a variety of issues have appeared especially with reference to economic and market (industry) concentration.²⁰

The factual inputs from Census data typically, however, are relatively crude. Undoubtedly, though, the availability of Census data has served to accentuate the industry type of analysis. The feasibility and high productivity of industry type researches perhaps explains the shift in Mason's position in his 1957 volume where he states:

When the term "market" is used, a Marshallian industry is meant; that is, a census industry, appropriately adjusted for product and spatial considerations. Unless we can use the conception of the market, and with it, properly rectified data, the field of Industrial Organization is a wilderness. Triffin, with his general equilibrium of the firm, has attempted to lead us up the garden path; and if we expect to retain our virtue, we had better retreat as rapidly as possible to the shelter of the Marshallian industry.²¹

Insofar as there is a common framework of analysis in industrial organization, it is the so-called "market structure-conduct-performance" approach. Although the causal relations presumably run from structure to conduct to performance, the interest and the analysis usually goes directly from structure to performance results or from structure to combined conduct-performance results in which business behavior (conduct) is inferred from performance results. Usually, three so-called "elements of structure" are used;

namely, size and size distribution, conditions of entry, and product differentiation. There is a wide variation as to what should be included from the other environmental factors and conditions affecting behavior and performance, ranging all the way from the above three elements to every conceivable objective aspect of the physical, economic, technological, and social universe that might conceivably affect decision making. But rarely does the analysis focus explicitly and sharply upon the internal organization and decision making of enterprises despite the explicit statement in the 1938 position paper, "firms are not, regardless of what economic theory may suppose, undifferentiated, profit-maximizing agencies which react to given market situations in ways which are independent of their organization . . . management . . . is influenced not only by market pressures, but also by considerations internal to the firm."²² This omission is crucial, since the center of interest is the large-scale corporation and oligopoly.

The most important issue for the field of industrial organization is how to bring the large diversified corporation within the framework of analysis. The crux of the matter is whether the market structure framework can be employed at all; in other words, is it relevant? If such large corporations are free of the market, as some allege, it would seem futile to try to analyze their behavior and performance results in a market structure framework. The focus of research then should be on internal organization, policies and strategies, and their performance results. Orientation should then be from performance results back into internal organization and decision making. But if there is a significant amount of market determinism and constraint, even if only for a period of time under given structural characteristics, it would seem reasonable to use the market structure framework of analysis. Such use would also be an important managerial tool for internal purposes as well as for public policy, especially in antitrust enforcement.

II. Summary of Views Expressed by Others

As a background for and check on my own judgment, I consulted with twenty-one persons active in the field of industrial organization, either as research contributors or as users of the research output, but omitting persons on this program, since they can speak for themselves. My inquiry was in terms of (1) the specific contributions of the field to date, (2) the chief weaknesses, and (3) sug-

(Yale Univ. Press, 1959); Jesse W. Markham, *Competition in the Rayon Industry* (Harvard Univ. Press, 1952); Jesse W. Markham, *The Fertilizer Industry: Study of an Imperfect Market* (Vanderbilt Univ. Press, 1958); J. W. McKie, *Tin Cans and Tin Plate* (Harvard Univ. Press, 1959); Walter J. Mead, *Competition and Oligopoly in the Douglas Fir Lumber Industry* (Univ. of California Press, 1966); Reed Moyer, *Competition in the Midwestern Coal Industry* (Harvard Univ. Press, 1964); Willard F. Mueller and Leon Garoian, *Changes in the Market Structure of Grocery Retailing* (Univ. of Wisconsin Press, 1961); William H. Nicholls, *Price Policies in the Cigarette Industry: A Study of "Concerted Action" and Its Social Control* (Vanderbilt Univ. Press, 1961); Merton J. Peck, *Competition in the Aluminum Industry* (Harvard Univ. Press, 1961); R. B. Tennant, *The American Cigarette Industry* (Yale Univ. Press, 1950). The Federal Trade Commission has published a number of industrial reports in the food, baking, antibiotics, and tire industries. The 1966 studies published under the auspices of the U. S. National Commission on Food Marketing focus upon organization and competition in sectors of the food industries.

²⁰ See N. R. Collins and Lee E. Preston, *Concentration and Price Cost Margins in Manufacturing Industries* (Univ. of California Press, 1968), for summaries of previous studies on concentration.

²¹ Mason, *op. cit.*, p. 5.

²² *Op. cit.*, p. 62.

gestions for the future. Time does not allow a full report on these statements of appraisal and of position. To some extent, of course, my own final observations will reflect what I learned, but not entirely so because of the wide diversity and conflicts of views and my own differences with some of them.

The diversity among these commentaries supports the view that industrial organization is not a clearly defined homogeneous entity. A few even held that industrial organization has no logic of its own as a separate field. The field continues to be broadly eclectic and evolving. It was generally agreed that the market structure, conduct, performance framework has been useful for empirical, theoretical, and public policy purposes. There was a heavy stress both on empirical work and on refinements of theoretical analysis. In the main, it was felt that the contributions to theoretical refinements and in empirical work have been substantial, but there was some dissent.

The criticisms stress the limitations of data, particularly the heavy reliance on Census data, together with limitations of measurement at all levels, the lack of normative decision models, and the lack of development of norms in general, the lack of use of refined game theoretical models, the relatively undue interest in homogeneous oligopoly models, the lack of attention to vertical market structures, the failure to develop an integrated theory or complex of theories of oligopoly, and the lack of a formal language, on the one hand, and conversely the relatively weak institutional footing on the other, among other things. Clearly, these are very sharp differences of judgment and attitude.

As for the future, the views ranged all the way from the pessimistic judgment that the field is falling apart through the moderate view that we need more of the same, only better, to a variety of suggestions for theoretical and empirical advances. In fact, sixty-two specific suggestions were made—a spectrum much too wide and diverse to present and interpret here except with broad brush strokes. A few suggested that the concept of market structure should be reexamined and reformulated in the perspective of developments in other disciplines, including general systems analysis. There was general agreement that advances must continue on both the empirical and theoretical fronts and in the application to public policies, especially in antitrust and in public regulation in general. The high need for theoretical-empirical work in the field of oligopoly and especially on problems of diversification and conglomerateness

was stressed. But there was no agreement as to the most effective approach. There was considerable stress on the further development of the theory of the firm, including growth theory, with emphasis on oligopoly theory and especially heterogeneous oligopoly and diversification. The hope was expressed that a wider variety of specific industry and cross-industry studies would appear together with comparative international studies. A few emphasized the need for relating the internal organization goals and decision making of corporations to market structure factors, especially when the internal allocation of resources has replaced the market. In the main, however, the emphasis was on external factors and upon the industry. Although the replies were not explicit, the general tenor indicated confidence in the received industry approach; but there was a minority view that the individual enterprise should be the focal point for entering into the broader industry and economic system relationships, especially in the case of multiactivity-multimarket firms. There was some difference as to the relative reliance upon general theory or intensive case studies in which business facts, economic theory, and legal theory can be correlated.

III. Prospect

The public policy interest and applications are now propelling the field of industrial organization so strongly that it would be impossible to stop the momentum. Consequently, professional workers in this field have a high responsibility and an unusual opportunity. The market structure-conduct-performance approach is now a basis of analysis and for judgments in much of the work in the antitrust field in both the Department of Justice and the Federal Trade Commission. The selection and analysis of cases is being influenced to some extent by this framework of analysis as is much of the basic research and data collection. The merger guidelines issued in 1968 by the Department of Justice are based on selective market structure criteria. And of even greater significance, courts—and especially the United States Supreme Court—are drawing heavily upon some of the hypotheses, research results, and generalizations of the literature. This is especially true of merger actions under the 1950 Celler-Kefauver Act. But the orbit of actual and potential application is much wider, and it must be so if our national economic policy of competition is enforced.

In the late 1950's, the courts—especially the United States Supreme Court—were requesting more and better economic evidence and analysis

than was being presented to them.²³ So far as I can judge, this is still true. I have been rechecking the opinions and briefs in antitrust cases presented to the United States Supreme Court beginning with *Brown Shoe* in 1962. Economic analysis—and especially the market structure framework of analysis, and more particularly the evidence and interpretations of concentration, market occupancy (share), oligopoly, entry, and product differentiation are being increasingly invoked. The strategic cases are *Brown Shoe* (1962), *Philadelphia National Bank* (1963), and *Procter and Gamble-Clorox* (1967) in terms of economic analysis.

Currently, high interest in and policy proposals with respect to the conglomerates (so-called) have focused attention on the most complex issue of antitrust enforcement. In the background are, of course, the reports of the P. C. Neal²⁴ and the George J. Stigler²⁵ task forces. Things are moving rapidly, and much of what scholars may consider relatively unjelled or incomplete or inadequate has been drawn into the public arena.

The field of industrial organization has been acting as a bridge between economic theory and public policy but on a selective basis. But the orbit of influence is widening. For example, there is now a more systematic working relationship and application in agriculture partly because the focus of interest increasingly is on the processing industries.²⁶ It is of high importance that a body of scholars with a basic interest in public policies and a footing in the economic theories of competition and a strong empirical interest continue in recognizable association regardless of whether the trade name be industrial organization or a more appropriate cognomen.

The general case, in terms of microtheory, really is oligopoly, even under the large numbers of monopolistic competition, when allowance is made for spatial and product and service differentiation and market segmentation. But the important issue in public policy is not the presence of a horde of local oligopolists, but the large corporation. From this standpoint, Edward S. Mason's instincts were sound in 1938, when he held that analysis must be from the vantage point

of the individual market participant. This is not to say, at all, that specific industry and cross-industry studies should not continue to be made. This line of research has been and will continue to be productive.

But the modern, powerful, diversified corporation cannot be fitted neatly into the partial equilibrium framework. But this is not to say that market structure analysis is not applicable to it. A diversified corporation is still a collection of products, product lines, and geographical divisions, and so on, in each of which market structure analysis may be applicable. It is highly important that analysis be focused, at least to begin with, in terms of the position of the enterprise in its various individual markets and submarkets at a given time. It is most interesting and perhaps significant that Edward H. Chamberlin, too, in his illuminating paper, "Monopolistic Competition Revisited," published in 1951, took the same position. Chamberlin came to view the economic system as a network of interrelated firms, in which oligopolistic relations have much greater force than he had envisaged in his original formulation.²⁷

In a sense, I am suggesting that we should walk up Triffin's garden path²⁸ which Mason warned against in his 1957 volume. But the path need not lead into an uncharted wilderness because time and again, and perhaps, most frequently, the analysis of products and product lines and so forth could be in relation to recognized natural industry categories, data, and the appropriate norms of appraisal. Thus, a pure conglomerate, if such there be, would break up into a series of discrete units by definition, each subject to market structure analysis. Anything of interest beyond this would be related most likely to short-run financial or stock market motivation and objectives.

A most complex issue is the nature of the patterns and results of diversification by internal growth or by acquisition. Evolving patterns of diversification are, however, subject to theoretical analysis and interpretation. At a given period of time, market structure analysis will be able to relate the internal product unit to either a recognized industry grouping or to some other com-

²³ E. T. Grether, "Economic Analysis in Anti-Trust Enforcement," *The Antitrust Bull.*, Jan.-Feb., 1959.

²⁴ *Antitrust Law and Economics Rev.*, Winter, 1968-69, pp. 11-53.

²⁵ *Congressional Record*, June 16, 1969, pp. 56472-82.

²⁶ Cf. R. L. Clodius and W. F. Mueller, "Market Structure Analysis as an Orientation for Research in Agricultural Economics," *J. of Farm Econ.*, Aug., 1961, pp. 513-53.

²⁷ E. H. Chamberlin, "Monopolistic Competition Revisited," *Economica*, Nov., 1951.

²⁸ For a recent appraisal of the Triffin position and a suggested approach, see R. E. Kuenne, "Quality Space, Interproduct Competition, and General Equilibrium Theory," pp. 225-32, in *Monopolistic Competition Theory: Studies in Impact*, ed. by Kuenne (Wiley, 1967).

petitive grouping, when the traditional industry-relevant market type of analysis is inappropriate.²⁹

The more interesting, more complex and more important, and more exciting issues arise out of (1) the synergetic relations among the internal product lines, subsidiaries, and so forth, of the large diversified corporation and (2) the continuing interactions between internal firm organization policies and practices and market structures. Careful detailed studies are needed of individual corporations, in which internal organization and policies at a given time and over periods of time are related to market structures and structural changes with the clear awareness that the successful, well-managed corporation will use its discretionary power to optimize both synergetic effects and favorable market structure factors. We know too little about the nature and effects of the interdependencies among the product lines and other units of activity assembled under one banner.

Combined market structure-internal organizational-action parameter type of analysis could become, and, to some degree probably is now, a major planning tool for large diversified corporations in the American environment. This approach combines market structure analysis as developed in this country with action parameter type of analysis which has reached a higher level of refinement in the Scandinavian countries.³⁰ Scandinavian scholars have been more successful relatively in bringing individual firm behavioral analysis into relation with the framework of economic theory than in this country.

It is of utmost importance to investigate and interpret competitive behavior on the numerous fronts in which a large diversified enterprise is engaged. Such research, in my opinion, is feasible and could be planned to show both short-period and longer-period results. It would also take several strides forward in bringing the economic theories of the firm and the organizational and behavioral theories into better working relationships. It is amazing and certainly revela-

tory as to lack of integration of scholarship in the United States that a 1965 mammoth collection of essays on organizations³¹ does not even index the field of industrial organization. The same observation applies to an antecedent volume on the behavioral theory of the firm.³²

Looking even further into the horizons, perhaps the numerous and growing body of scholars in management science, operations research, game theory, and their subspecialties could come to develop more affinity with economic theory and vice versa. Much of the work in business strategies, including game theoretic models, might help to compensate for the inherent negative stance of most oligopoly models of economic theory. Similarly, too, much of the analysis in terms of business strategies tends to be so narrowly limited as to contribute little to the framework of general theory except, perhaps, as it relates to some formulation of systems analysis. The general may well be lost in a host of endeavors to solve particular problems or in investigations narrowly confined so that specific answers can be obtained. But in this respect, we are probably no worse off than other social sciences and the life sciences. Many areas of learning seem to be running over with specific, often minute, collections of data and highly specific, narrow investigations with too little effort at integration.

The field of industrial organization was pushed forward appreciably by the opportunities and motivation for systematic cross-disciplinary interaction provided by the Merrill Foundation at Harvard over a period of years in the early 1950's.³³ Two of the book-length products stimulated by this group interaction in themselves represent a high return on the investment of time and other resources; namely, J. S. Bain, *Barriers to New Competition*, and the Kaysen and Turner volume on *Antitrust Policy*.

There is high need for another, but wider and more continuous, interdisciplinary endeavor. I made such a recommendation before the Joint Economic Committee of Congress on September 23, 1959. The need and the opportunity and likelihood of substantial success are much greater now. It should be possible now at least to get agreement on the conceptualization and definition

²⁹ Cf. C. D. Edwards, "The Changing Dimensions of Business Power," in *Das Unternehmen in Der Rechtsordnung*, 1967, pp. 237-60.

³⁰ For the original development of the action parameter concept, see Ragnar Frisch, "Monopoly-Polypoly—The Concept of Force in the Economy," in *Int. Econ. Papers*, No. 1, translations prepared for the International Economic Asso. (London: Macmillan & Company, 1951). The original article appeared in April, 1933, as supplement to *Nationalökonomisk Tidskrift*. For a general summary, see G. Mickwitz, *Marketing and Competition* (*Societas Scientiarum Fennica*, Helsingfors, 1959).

³¹ J. G. March, ed., *Handbook of Organizations* (Rand McNally, 1965). The writings of only a few scholars in industrial organization are even cited.

³² R. M. Cyert and J. G. March, *A Behavioral Theory of the Firm* (Prentice-Hall, 1963).

³³ For details, see Mason's preface to Carl Kaysen and D. F. Turner, *Antitrust Policy: An Economic and Legal Analysis* (Harvard Univ. Press, 1959).

of market structure. Too frequently, vertical market structure relations are overlooked entirely or not brought systematically into the analysis. One of the finest contributions would be to push the investigation of the interaction of structural and behavioral factors into long-period analysis so that we would have a better understanding of the long-run determination of market structures. The earlier ambitions for the development of systematic classifications based on market structures which were given up might be revived with the aid of the computer. An enormous and growing body of fragmented evidence is waiting for more systematic interpretation. The hypotheses and the results of the testing should be gathered together into a reasonable framework. It would be possible to improve the tools, framework, and empirical resources for the conjoined analysis of oligopoly and diversification. Certainly, quantitative analysis including econometric-type measurements could be advanced; relatively, industrial organization has lagged on this score. Hopefully, the highly varied, numerous investigations in management science, etc., oriented in terms of the internal organization and decision making of enterprises, could be brought into systematic relation with the more objective,

external-market structure-environmental types of research. Then, too, norms of performance could be related more reasonably to organizational form and structural and behavioral variables.

In the biological sciences, in which specialization and subspecialization apparently are as great or even more so than in economics and the social sciences, the field of immunology, because of its direct bearing on public health, provides both a bridge among the specialties and rays of hope in the search for general understanding and application—and also access to sizable resources for research.³⁴

Market structure analysis is a systematic means of analyzing the linkages of enterprises into the broader macrorelations of the competitive market system and of the economy. It could be a major means for broader integration and understanding. Undoubtedly, substantial resources can be obtained from both governmental and private sources for research because of the strategic significance both for public and private policy and decision making.

³⁴ Paul A. Weiss, "Living Nature and the Knowledge Gap," *Saturday Rev.*, Nov. 29, 1969.

INDUSTRIAL ORGANIZATION: PRICE MODELS AND PUBLIC POLICY

By EUGENE M. SINGER
New York, New York

A little over a decade ago the collected works of Dean Edward S. Mason, of the Harvard Graduate School of Public Administration, were published in a volume entitled, *Economic Concentration and the Monopoly Problem*. In the introduction [16, p. 4] Dean Mason noted: "No one who is other than eclectic, methodologically speaking, has any business in the field of business organization. The price theorist is, sooner or later, sure to become lonesome in the exclusive company of his market models. The collector and classifier of census data and price statistics will, at some stage, begin to wonder about the relevance of his aggregates and averages."

Many of the transitions from the domains of price theory to Census data are becoming important in the statements issued by officials in government agencies and courts of law [15]. When spoken by an attorney general, written in a decision by the Supreme Court, or released in a committee report with the seal of the President of the United States, economic data may assume an aura of fact. Conclusions of the modelmaker tend to be adopted unencumbered by the assumptions. Preliminary studies of a limited number of industries become foundations for positive statements concerning the whole manufacturing sector. As the tide runs with the ideological turbulence of politics, the theorist and Census data classifier may find the crossing increasingly difficult.

The use of an economic model in industrial organization need not depend on its ability to define a specific course of action for a policy-maker. Within the framework of *ceteris paribus* assumptions, a model can suggest almost any alternative with regard to price or output depending upon which terms are treated as variables or constants. Therefore, the phrase "economic theory suggests" can generally be used by either side of any argument in industrial organization. Unfortunately, when theory is employed to buttress, by way of confirmation, a finding based on a sparse set of data, there is an inference that the theoretical model predicts a single outcome—which conveniently coincides with the conclusion sought to be drawn from that data.

In this context the static analysis of a model such as pure monopoly is used as a foundation for a structural approach to industrial organization. The burdensome problems of complementarity in

demand functions, interdependence between supply functions, and externalities, are side-stepped. The replacement is a simplistic form of analysis relating one variable, such as a concentration ratio, which is identified as an element of structure, with another variable, such as profits, which is identified as a performance element. The terminology of "structure" and "performance" does not avoid the classic problems raised in economic models. Nor can it lessen the analytical value of models and their useful framework for studying variables. The structural terminology, to the extent that it permits an oversimplification of the processes of competition, may be more of a retrograde than progressive step in the field of industrial organization.

I. Trends in Concentration

Structural approaches in antitrust policy have often been related to a need to slow or reverse an assumed rapidly increasing or high level of concentration. In its first case under the 1950 Celler-Kefauver Anti-Merger Act, amending Section 7 of the Clayton Act, the Supreme Court stated in the Brown Shoe decision [5, pp. 315-16]: "The dominant theme pervading congressional consideration of the 1950 amendment was a fear of what was considered to be a rising tide of economic concentration in the American economy." Similarly, in June, 1969, the Attorney General of the United States observed: "The danger that this superconcentration poses to our economic, political and social structure cannot be overestimated" [17].

Considerable doubt exists as to whether there has been any recent substantial increase in economic concentration. In his review of the literature on economic concentration up to the mid-1950's, Dean Mason concluded: "I think it has been demonstrated that between 1931 and 1947 there was no substantial increase in concentration in manufacture, and at least serious doubts have been cast on the existence of any such tendency over the last fifty years" [16, p. 43].

A recent study of trends in economic concentration in the United States economy was released in January, 1969, by the Cabinet Committee on Price Stability established by President Johnson [24]. Study Paper No. 2 concluded that "average

market concentration of manufacturing industries has shown no marked tendency to increase or decrease between 1947 and 1966, according to an analysis of 213 essentially comparable industries" [24, p. 58]. Consumer and producer goods industries were found to exhibit contrasting patterns: "The average decline in concentration in producer goods industries was offset by a substantial upward movement in consumer goods industries" [24, p. 60]. This upward movement was found by dividing all consumer goods into three categories: undifferentiated, moderately differentiated, and highly differentiated. In the nineteen-year period under consideration, the undifferentiated consumer goods showed an increase in average concentration of less than 1 percentage point, the moderately differentiated goods showed an increase of 4.5 percentage points, and the highly differentiated goods had an increase of 12 percentage points. Thus, the crux of the Price Stability Committee's conclusion of a "sharply upward movement" in consumer goods industry concentration lies within the highly differentiated sector.

Table 1 shows the list of seventeen highly differentiated consumer goods industries used by the Committee.¹ Throughout their study the Committee used unweighted averages of concentration ratios. Thus, the concentration ratio of the top four companies producing Chewing Gum is given identical weight in being averaged with the ratio of the leading producers of Motor Vehicles and Parts. By definition, each of these industry concentration ratios is a percent with a different base; the unweighted averaging of the percentages is therefore statistically erroneous [9, p. 135]. In Table 1 the Cabinet Committee's list of highly differentiated consumer goods industries is, con-

trary to their procedure, weighted by value of shipments. Two of the industries, Wines and Brandy and Motor Vehicles and Parts have concentration ratios based on value added rather than value of shipments, due to the presence of extensive duplication. These industries have been placed in Part II of Table 1 along with Greeting Cards, which is apparently misclassified as highly differentiated. If these three industries are eliminated from the group of seventeen industries, there is an unchanged weighted average concentration ratio of 48 in 1947 and also in 1963. Only by including in the weighting a questionable value of shipment figure for Motor Vehicles and Parts can one produce an upward movement. Almost all of the increase in concentration in this industry occurred by 1954 when the top four companies accounted for 75 percent of the industry value added. Since 1954 there has been a significant increase in the relative number of imported cars; however, these and other imports are excluded in the computations of concentration ratios in Table 1.

The important point is that public policy conclusions cannot be drawn meaningfully with respect to all highly differentiated consumer goods from a list of so few industries that the inclusion or exclusion of one industry from a weighted average causes widely differing results.

II. Concentration and Profits

A second justification for a simplistic structural approach to public policy rests on the proposition that any increase in concentration is directly identified with a lessening of competition. The process of competition requires a complex analysis of manufacturing and distribution techniques, pricing, innovation, and multiple-product relationships. In contrast, the term "concentration" requires only an arithmetic computation of a percentage. By substituting "concentration" for "competition" in the analysis of antitrust problems one can conveniently solve the problem of ascertaining a tendency for a substantial lessening of competition by finding an increase in concentration. This philosophy was expounded by the Price Stability Committee's Study Paper No. 2: "Market concentration is directly related to the intensity of competition in an industry" [24, p. 54] (compare to [25]).

A significant part of the support for this proposition is based on studies which show a correlation between concentration and profit rates.² Profit

¹ The method by which the seventeen industries were selected was footnoted in the study [24, p. 60] as follows: "Generally speaking, industries classified as undifferentiated make advertising expenditures of less than 1 percent of sales, and those classified as highly differentiated make substantial expenditures for advertising, often in excess of 10 percent of sales and usually are heavy users of television advertising media. See Federal Trade Commission, *Industry Classification and Concentration* (1967)." Inspection of the FTC document, which is the basis for the differentiation classifications, shows the identical footnote plus the following sentence: "Incomplete information makes classification into differentiation categories somewhat subjective" [10, p. 18]. The result is that an industry can be considered highly differentiated in Table 1 even though its advertising expenditures are less than 1 percent of sales. The range of subjectiveness is apparent when an industry such as greeting cards is included in Table 1. It is difficult to perceive the distinguishing characteristics that would enable a consumer to differentiate between the cards of different manufacturers.

² Many of the leading studies are footnoted by Professor Brozen [6, p. 125, fn. 16].

TABLE 1
LIST OF HIGHLY DIFFERENTIATED CONSUMER INDUSTRIES USED BY
PRESIDENT JOHNSON'S CABINET COMMITTEE ON PRICE STABILITY

SIC Code	Industry	Top Four Company Shipments Concentration Ratio		Industry Value of Shipments in Thousands of Dollars	
		1947 (1)	1963 (2)	1947 (3)	1963 (4)
	<i>Part I:</i>				
2043	Cereal Preparations.....	79	86	284,320	625,058
2072	Chocolate and Cocoa Products....	68	75	349,907	479,085
2073	Chewing Gum.....	70	90	148,286	224,844
2082	Malt Liquor*.....	21	34	1,316,005	2,315,068
2085	Distilled Liquor*.....	75	58	870,235	1,090,462
2086	Bottled and Canned Soft Drinks...	10	12	748,196	2,210,920
2087	Flavorings.....	50	62	313,573	729,704
2111	Cigarettes*.....	90	80	1,131,891	2,655,346
2121	Cigars*.....	41	59	311,401	358,463
2834	Pharmaceutical Preparations.....	28	22	941,290	3,314,323
2844	Toilet Preparations.....	24	38	381,376	1,792,662
3421	Cutlery (Includes Razor Blades)...	41	66	142,571	283,352
3633	Household Laundry Equipment....	40	78	442,297	760,161
3861	Photographic Equipment.....	61	63	457,497	1,851,213
	Subtotal.....	698	823	7,838,845	18,690,661
	<i>Part II:</i>				
2084	Wines and Brandy†.....	26	44	167,418	368,281
3717	Motor Vehicles and Parts†.....	56	79	11,564,913	36,181,007
2771	Greeting Cards.....	39	57	123,856	345,961
	Subtotal.....	121	180	11,856,187	36,895,249
	Total.....	819	1003	19,695,032	55,585,910
	Part I Weighted Concentration Ratio.....	48	48		
	Parts I and II Weighted Average Concentration Ratio.....	52	68		
	Parts I and II Unweighted Average Concentration Ratio (Price Stability Committee).....	48	59		

* Value of production data.

† The presence of substantial duplication in the shipment figures for SIC 2084 and 3717 resulted in the Census Bureau computing the concentration ratios of these industries in terms of value added. The shipment figures for these industries are shown, with the duplication, in columns (3) and (4). If value added is used as a weight instead of value of shipments, the weighted average in 1963 of Part I is 47, and of Parts I and II, 64. Census concentration ratios in terms of value added are not available for 1947.

SOURCE: [10] [20] [24, p. 58] [26] [27].

rates are used as an index of competition with the inference that competition will be lessened as profit rates increase. To quote the final report of President Johnson's Council of Economic Advisers: "Numerous studies have shown a significant relationship between high concentration and

high profit rates—an indication of weak competitive pressures" [22, p. 162].

Such a contention implicitly focuses on a static price model with a monopolist having the power to adjust price and faced with constant demand and cost factors. If concentrated industries lower

costs by improving manufacturing processes or expand sales by introducing new and improved products, their increased profits are still assumed to be the result of monopoly power under the above proposition. Profits flowing from these forms of beneficial economic activity should be encouraged; instead, profits are stigmatized with the odium of monopoly power. Similarly, if the leading companies in concentrated industries were to invest more heavily in capital equipment than labor intensive smaller firms, their profits would be deemed the results of barriers to entry. The encouragement of internal growth through capital expenditures may not only be the very activity we should seek to stimulate but may account for precisely the higher profits being earned.³

Two often cited studies which consider the relationship between concentration and profit rates are Joe Bain, *Barriers to New Competition* [2], which covers the periods 1936 to 1940 and 1947 to 1951, and an article by Michael Mann [13], covering the period 1950-60.⁴ On an examination of twenty industries, Bain concludes that "there is evidently a clear association between the rate of profit earned in an industry and the degree of seller concentration in it" [2, p. 195]. Mann examines thirty industries, which include seventeen listed in the Bain study, and concludes: "One clear outcome of Bain's work and of this study is that the 'monopoly problem' appears to exist most noticeably in those industries which are highly concentrated and have high barriers to entry. If public policy seeks to improve resource allocation, industries with these structural characteristics seem to be a good place to start. . . . It is these industries which might be called to the attention of antitrust authorities" [13, p. 300].

In both the Bain and Mann studies there is no weighting of either the profit rates of different sized firms within an industry or the profit rates

of different industries. The industry average rates of return (net income to average net worth) are not weighted averages for all firms classified in the industry but are unweighted averages of the profit rates of a few leading firms.⁵ In the Bain study of twenty industries there are eight industries which have only two firms included in the computation of the industry average profit rate. In the Mann study of thirty industries more than half of the industry profit rates are derived from only two or three firms in each industry. The profit rates for these firms cover all activities, although an attempt was made to select firms which received at least 50 percent of their sales revenue from the product of the industry in which they were placed. The capital structures differ considerably between firms within an industry and among industries. For example, Mann adjusts downward the aluminum industry profit rate because, "these companies have very high debt to equity ratios which inflates their rates of return on net worth considerably." No upward adjustments in profit rates were deemed necessary in the more labor intensive, lower concentrated industries which had lower debt-equity ratios.

The industries selected in these studies are not all taken from the manufacturing sector and do not have a common degree of product classification in terms of the Standard Industrial Classification Code (SIC). The Bain study includes an industry entitled, "Shoes (high priced men's)," for which there is not even a 7-digit SIC Code [5]. The Mann study includes three industries (Sulphur, Nickel and Bituminous Coal) which are not included in the manufacturing sector, one 2-digit industry (Textile Mill Products), and one 7-digit industry (Shoe Machinery). Both studies include a 3-digit industry (Farm Machinery and Tractors) which is a combination of two 4-digit industries. The remaining industries in these studies are from the manufacturing sector and appear to be at the 4-digit classification level.

In Table 2 the twenty industries used in the Bain study for the period 1947-51 are listed along with their nearest 4-digit SIC Code. The industries were divided by Bain, as shown in column (6), into two groups: industries in which the top eight company concentration ratios in 1947 are above 70 percent, and those under 70 percent.⁶

³ Price-cost margins which include in the gap depreciation in addition to profits would be expected to be larger for the more capital intensive firms associated with concentrated industries. A finding that concentration alone rarely explained as much as a fourth of the variations in price-cost margins among 4-digit industries [8, p. 108] is further attenuated by the fact that variations in the margins can be the result of depreciation rather than profits. Furthermore, concentration ratios may serve as a proxy variable for average firm size in an industry [11].

⁴ For example, both are cited in the *Cabinet Committee Study* [24, p. 100, fn. 7 and 8]. The Bain study has been cited by the Supreme Court [28]. The Mann study was one of the selections in the first issue of *The Journal of Reprints for Antitrust Law and Economics*, Summer, 1969.

⁵ Professor Bain states: "The data are hardly complete enough to justify the arduous task of weighting individual observations" [2, p. 195]. However, the unweighted averaging of percentages representing rates of return which have different bases is statistically erroneous [9, p. 135].

⁶ Professor Bain does not specify whether his concentration ratios are computed in terms of value of ship-

TABLE 2
LIST OF TWENTY INDUSTRIES USED IN THE BAIN STUDY OF BARRIERS TO NEW COMPETITION

1947 Nearest 4-digit SIC Code (1)	Bain Industry Designation (2)	Bain Unweighted Average Profit Rate of Leading Firms, 1947-51 (3)	1947 Industry Value Added, Unadjusted (4)	1950 Top Four Company Employment Concentration Ratio (5)	Bain's Top Eight Company Concentration Classification: (I) Over 70% (II) Under 70% (6)
Part I. Manufacturing Industries with Top Four Company Employment Concentration Ratio Above 70%					
2085	Liquor.....	18.6	472,357	74	I
2111	Cigarettes.....	12.6	368,417	81	I
2825	Rayon.....	18.0	433,285	76	I
3011	Tires and Tubes.....	12.7	650,165	78	I
3272	Gypsum Products.....	15.4	67,879	90	I
3331	Copper.....	14.6	93,196	91	I
3411	Metal Containers.....	10.7	231,503	77	I
3572	Typewriters.....	18.0	121,385	90	I
Part II. Manufacturing Industries with Top Four Company Employment Concentration Ratio Below 70%					
2011	Meat Packing.....	5.1	977,144	47	II
2033	Canned Fruits and Vegetables..	9.8	608,768	20	II
2041	Flour.....	10.1	410,774	26	II
2841	Soap.....	15.8	499,792	69	I
2911	Petroleum Refining.....	12.9	1,494,474	39*	II
3141	Shoes.....	11.0	751,528	22	II
3312	Steel.....	11.2	2,601,582	54*	II
3241	Cement.....	14.3	247,936	29	II
3717	Automobiles.....	23.9	3,544,924	59	I
3951	Fountain Pens.....	21.8	89,897	48*	I
Part III. Other Industries Used in Bain Study					
— 3-digit	Shoes (high-priced men's).....	13.4	—	—	I
	Farm Machinery and Tractors..	13.4	—	—	II
	WEIGHTED AVERAGE RATE OF PROFIT				
	Part I.....	15.0			
	Part II.....	15.1			
	UNWEIGHTED AVERAGE RATE OF PROFIT				
	Part I (Bain).....	16.6			
	Part II (Bain).....	11.0			

* 1954 employment data used; 1950 data unavailable.

Note: the 1947 SIC 4-digit titles for Part I are Distilled Liquor, Except Brandy; Cigarettes; Synthetic Fibers; Tires and Inner Tubes; Gypsum Products; Primary Copper; Tin Cans and Other Tinware; and Typewriters. The SIC titles for Part II are Meat Packing, Wholesale; Canning and Preserving, Except Fish; Flour and Meal; Soap and Glycerin; Petroleum Refining; Footwear, Except Rubber; Steel Works and Rolling Mills; Cement, Hydraulic; Motor Vehicles and Parts; and Pens and Mechanical Pencils.

SOURCES: [2, pp. 195-97] [20] [26].

Table 2 reclassifies these industries into Parts I and II depending on whether their top four company employment concentration ratio in 1950 is above or below 70 percent.⁷ The industries, Farm Machinery and Tractors and Shoes (high priced men's) are relegated to Part III which excludes these industries from further study. The former is a 3-digit industry and the latter already has its nearest 4-digit industry included in Part II as Shoes. If these eighteen industries are weighted by industry value-added figures for 1947, a conclusion opposite to Bain's is reached. The unweighted average of Bain under his method of classifying industries showed the concentrated industries earning an average of 16.6 percent and the less concentrated industries earning 11.0 percent.⁸ In Table 2 the weighted average profit ratios of Part I, the concentrated industries with top four company employment concentration ratios over 70 percent, had an insignificantly

lower average profit rate (15.0) than the less concentrated industries in Part II (15.1).⁹

The data in the Mann study can also be rearranged, within the boundaries of reasonable discretion, to produce opposite results. In Table 3 the same method of reclassifying industries according to top four company employment concentration ratios is followed. The weights are value added, adjusted, for 1954. Six of Mann's thirty industries are relegated to Part III: the three nonmanufacturing industries, a 2-digit and a 3-digit industry, which are too broad for inclusion, and a 7-digit industry which is deemed too narrow. Thus, the remaining twenty-four industries in the Mann study are from the manufacturing sector and have 4-digit SIC codes. The unweighted average profit rate of Mann showed his concentrated industries with 13.3 percent and the less concentrated industries with 9.0 percent.¹⁰ In Table 3 the weighted average profit ratio of Part I, the concentrated industries with top four company employment concentration ratios over 70 percent, again were found to have a lower average profit rate (12.0) than the less concentrated industries in Part II (12.2).¹¹

ments, value added, or employment data. Bain states that he draws "upon Census national seller concentration data for 1935 and 1947 in most cases, and on generally equivalent data (for copper, automobiles, and quality fountain pens, for example) where Census data will not suffice" [2, p. 196].

⁷ In Table 2 employment concentration ratios were not available in 1947. These ratios were available in 1950 for all but three industries. The 1954 employment concentration ratios were used for the latter three industries. Value-added concentration ratios were not available for seventeen of the industries in 1947. Value of shipment concentration ratios were not available in 1947 or 1954 for six industries: SIC 2085, 3331, 2011, 2033, 3312 and 3717.

⁸ Professor Bain, in contrast to his earlier study of forty-two industries [3], did not perform any statistical tests between the means of the average profits of the two groups of industries. He noted that "the data are incomplete in a number of cases, and not of such a character as to justify extensive treatment by formal statistical techniques" [2, p. 195]. Nevertheless, Professor Mann subsequently tested the significance of the difference in average profit rates of the twelve industries in Bain's unconcentrated group and the eight industries in the unconcentrated group. Mann concluded: "The difference in average profitability between those industries with concentration ratios above 70 percent for the top eight sellers and those with concentration ratios below 70 percent is statistically significant at the .01 level" [14, p. 393]. On this basis, Mann subsequently asserted without qualification as to the limited number of industries: "Joe Bain's finding that concentration influenced profitability in the years 1947-51" [14, p. 393]. The use of statistical significance tests may be questionable in studies of a small number of industries. The group of industries chosen seldom represents a "sample" with homogeneous characteristics which is randomly selected. Assumptions that a population, or at least the means of samples from a given population, is normally distributed, and that a linear relationship can be expected, are often open to doubt [1].

⁹ The weights employed are industry value-added figures. If 1947 concentration ratios in terms of value added and value of production data are used to supplement value of shipment concentration data where the latter are not available, the classification of one industry is changed (SIC 2841) from under to over 70 percent. Another industry (SIC 3331) has concentration data only in terms of employment. Recomputation of Table 2 with SIC 2841 classified over 70 percent in Part I and SIC 3331 remaining in Part II shows the weighted average rate of profit for Part I the same (15.1) as for Part II. In contrast, Table 2 can be recomputed using Bain's top eight company concentration classification shown in column (6) and value added in 1947 as weights. The weighted average rate of profit for Part I would then be nearly twice (20.0) that of Part II (10.6).

¹⁰ On the basis of twenty-one industries in a concentrated group and only nine industries in an unconcentrated group, Mann was able to conclude that "the difference between average rates of return is statistically significant at the .01 level" [13, p. 298].

¹¹ The weights employed are industry value-added figures. If 1954 concentration ratios in terms of value added and value of production concentration data are used to supplement value of shipment concentration data where the latter are not available, the classification of two industries is changed (SIC 2051 and 3717). Another industry (SIC 3331) has no available concentration data for 1954 other than employment figures. Recomputation of Table 3 with SIC 2051 (Bread and Related Products) and 3717 (Motor Vehicles and Parts) reclassified over 70 percent in Part I, shows the weighted average rate of profit for Part I (13.9) higher than Part II (11.0). The result illustrates as in Table 1 the substantial effect of Motor Vehicles and Parts in a weighted average of a limited number of industries. If Table 3 is recomputed using Mann's top eight company concentration classification shown in column (6) and value

TABLE 3

LIST OF THIRTY INDUSTRIES USED IN THE MANN STUDY OF SELLER CONCENTRATION AND RATES OF RETURN

1954 Nearest 4-Digit SIC Code	Mann Industry Designation	Mann Unweighted Average Profit Rate of Leading Firms 1950-60	1954 Industry Value Added, Adjusted	1954 Top Four Company Employment Concentration Ratio	Mann's Top Eight Company Concentration Classification: (I) Over 70%; (II) Under 70%
(1)	(2)	(3)	(4)	(5)	(6)
Part I. Manufacturing Industries with Top Four Company Employment Concentration Ratios Above 70%					
2073	Chewing Gum.....	17.5	92,227	72	I
2111	Cigarettes.....	11.6	676,593	75	I
2825	Rayon.....	8.5	698,680	73	I
2841	Soap.....	13.3	521,413	71	I
3011	Tires and Tubes.....	13.2	877,771	80	I
3211	Flat Glass.....	18.8	247,175	85	I
3272	Gypsum Products.....	14.4	165,271	86	I
3331	Copper.....	11.5	150,805	85*	I
3334	Aluminum Reduction.....	10.2	258,944	100	I
3411	Metal Containers.....	9.9	500,408	78	I
Part II. Manufacturing Industries with Top Four Company Employment Concentration Ratios Below 70%					
2011	Meat Packing.....	5.3	1,397,175	44	II
2033	Canned Fruits and Vegetables..	7.7	850,114	40	II
2041	Flour.....	8.6	330,634	32	II
2051	Baking.....	11.0	1,636,813	18	II
2052	Biscuits.....	11.4	420,132	60	I
2082	Beer.....	10.9	1,106,303	22	II
2085	Liquor.....	9.0	338,403	57	I
2834	Ethical Drugs.....	17.9	1,182,246	25	I
2911	Petroleum Refining.....	12.2	1,918,020	39	II
3141	Shoes.....	9.6	942,961	26	II
3221	Glass Containers.....	13.3	384,647	60	I
3241	Cement.....	15.7	725,000	32†	I
3312	Steel.....	10.8	4,674,286	54	I
3717	Automobiles.....	15.5	5,901,436	69	I
Part III. Other Industries Used in Mann Study					
none	Sulphur.....	21.6	—	—	I
none	Nickel.....	18.9	—	—	I
none	Bituminous Coal.....	8.8	—	—	II
2-digits	Textile Mill Products.....	6.9	—	—	II
3-digits	Farm Machinery and Tractors..	8.8	—	—	I
7-digits	Shoe Machinery.....	7.4	—	—	I
WEIGHTED AVERAGE RATE OF PROFIT					
	Part I.....	12.0			
	Part II.....	12.2			
UNWEIGHTED AVERAGE RATE OF PROFIT					
	Part I (Mann).....	13.3			
	Part II (Mann).....	9.0			

* 1951 employment data used. 1954 data unavailable.

† 1958 value-added data used; 1954 data unavailable. The geographic regionalism of the cement industry is noted by Mann in his concentration classification.

Note: The 1954 SIC 4-digit titles for Part I are Chewing Gum; Cigarettes; Synthetic Fibers; Soap and Glycerin; Tires and Inner Tubes; Flat Glass; Gypsum Products; Primary Copper; Primary Aluminum; and Tin Cans and Other Tin Ware. The SIC titles for Part II are Meat Packing Plants; Canned Fruits and Vegetables; Flour and Meal; Bread and Related Products; Biscuits and Crackers; Beer and Ale; Distilled Liquor, Except Brandy; Pharmaceutical Preparations; Petroleum Refining; Footwear, Except Rubber; Glass Containers; Cement, Hydraulic; Steel Works and Rolling Mills; and Motor Vehicles and Parts.

SOURCE: [13, p. 299] [21] [26].

These two arithmetical exercises in Tables 2 and 3 have no more "disproved" the propositions advanced by these studies than the work of these studies has verified the relationship.¹² Our interest is not as concerned with whether a relationship exists between concentration and profits as the method by which this relationship has been purportedly established in certain studies. To the extent that other studies have relied on these methods for establishing a relationship between profits and concentration, the strength of their findings might be reassessed. Hypotheses other than market power which could account for this positive relationship might also be considered.¹³ More important is the conclusion that a group of twenty or thirty industries and a structural comparison of two variables cannot serve as a basis for drawing major policy statements applicable to the whole manufacturing sector of the American economy.

III. Deconcentration

In May, 1969, a report on antitrust policy prepared by a Task Force appointed by President Johnson was released. *The Presidential Task Force Report on Antitrust Policy* [23] is significant in showing how far an elementary structural approach can be carried in public policy. The Task Force proposed a Concentrated Industries Act which would force deconcentration upon any industry in which four or fewer firms had an aggregate market share of 70 percent or more for approximately five years and industry sales exceeded \$500,000,000. Firms in these industries would be required to reduce their market shares within a period of four years to no more than 12 percent. A firm could resist dissolution or divestiture only if it could demonstrate affirmatively that substantial losses in economies of scale would result. The method by which economies of scale could be meaningfully measured was not considered.

The Task Force utilized a price model to substantiate their structural approach. The following quotation from the Task Force is critical for understanding the theoretical underpinning of the elementary structural approach: "In a market

with numerous firms, each having a small share, no single firm by its action alone can exert a significant influence over price and thus output will be carried to the point where each seller's marginal cost equals the market price. This level of output is optimal from the point of view of the economy as a whole" [23, pp. 1-4].

In the static price model as formulated by Pareto, it is assumed not only that there exists perfect mobility of resources within the industry under consideration but for all input and output industries in the whole economy [29, pp. 18-24]. In a period of full employment, an increase in output in one industry will be at the expense of other industries. Professor William Baumol has noted: "In popular discussions one tends to think that the larger the output of any industry, the better off is society, but it is easy to see that this can result in a misallocation of resources . . ." [4, p. 215]. If a particular industry is deconcentrated, some firms will be weakened in their bargaining ability with suppliers and customers. Their marginal costs for raw materials may increase or the prices received for their output may be lowered. Resources will be reallocated with stronger industries gaining at the expense of weaker industries. Deconcentration of some but not all industries may merely benefit a few industries at the expense of other industries with either a negative or no net benefit conferred on the economy.

Further reason for rejecting the notion that a price model of pure competition in a single industry is a suitable guide for a public policy directed toward the whole economy is offered by Professors Lipsey and Lancaster in their "Theory of Second Best." They contend that in a mixed economy with a number of regulated and unregulated industries not operating under conditions of perfect competition it may be undesirable in terms of economic welfare to restructure less than all industries. "To apply to only a small part of an economy welfare rules which would lead to a Paretian optimum if they were applied everywhere, may move the economy away from, not toward, a second best optimum position" [12, p. 17].¹⁴

added, adjusted, in 1954 as weights, the weighted average rate of profit for Part I (13.3) is greater than Part II (9.7). These figures are close to the unweighted results of Mann notwithstanding the exclusion of six of his thirty industries which were relegated to Part III.

¹² This statement appears to be in agreement with Professor Bain: "The preceding findings are of course based on data so incomplete and on a sample so small that no conclusive confirmation or disconfirmation of our theoretical hypothesis has been possible" [2, p. 20].

¹³ Professor Brozen [6, p. 126] has suggested a number of alternative hypotheses.

¹⁴ Professor Markham has pointed out to me that the underlying reasoning of the theory of second best can be found in J. M. Clark's classic paper, "Toward a Concept of Workable Competition" [7, p. 242] written in 1940: "If there are, for example, five conditions, all of which are essential to perfect competition, and the first is lacking in a given case, then it no longer follows that we are necessarily better off for the presence of any one of the other four. In the absence of the first, it is a priori quite possible that the second and third may become positive detriments; and a workably satisfactory result may depend on achieving some degree of 'imperfection' in these other two factors."

In basic terms the theory of second best takes account of the interdependence of industries in the economy. Once the marginal cost of one input in one regulated industry is upset in a general equilibrium system, the balance of other inputs in all of the unregulated industries will have to be adjusted away from the perfect competition condition of price equal to marginal cost. Thus, the theory of second best casts doubts on theoretical grounds whether further deconcentration in selected industries would offer economic benefits to our country in the form of a better allocation of resources. The theory rejects the proposition that the greater the number of industries which approach the conditions of pure competition, the greater the likelihood that an optimum allocation of resources will occur.

IV. Conclusion

Elementary structural approaches in antitrust policy have been predicated principally on two grounds: (1) a need to slow or reverse an assumed rapidly increasing trend or high level of concentration and (2) a purported relationship between an increase in concentration and a lessening of competition. There exists considerable doubt whether there has been any rapidly increasing trend of concentration. The purported relationship between concentration and competition appears to rest on a weak association between profit rates of leading firms and concentration ratios for a limited number of industries. A simple correlation between one structure element such as a concentration ratio, and a performance variable such as a profit rate will not provide a critical link or easy formula for antitrust policy.¹⁵ The terms structure and performance may be convenient for organizing data, but unlike economic models they offer no analytical content in and of themselves.¹⁶ An approach which examines in detail the nature of competition in specific industries gives a greater opportunity for appreciating interdependence among variables and, in my opinion, is more likely to bring us closer to an understanding of the foundations of American enterprise.

¹⁵ Professor Almarin Phillips has observed that in structural studies "by far the most important reason for cautious interpretation . . . arises from the problem of assigning a direction of causation to the findings" [19, p. 193]. A variable such as profits can serve both as an incentive and effect. Furthermore, structural variables tend to serve as proxies for other variables: The unobserved variables are causative to the observed [19, p. 193] [18, pp. 88-116].

¹⁶ This sentence was suggested to me by Dr. Betty Bock, of the National Industrial Conference Board.

REFERENCES

1. P. ASCH, "Industry Structure and Performance: Some Empirical Evidence," *Rev. of Social Economy*, Sept., 1967, pp. 167-82.
2. J. BAIN, *Barriers to New Competition* (Cambridge, Mass., 1956).
3. ———, "Relation of Profit Rate to Industry Concentration: American Manufacturing, 1936-1940," *Q.J.E.*, Aug., 1951, pp. 293-324.
4. W. J. BAUMOL, *Economic Theory and Operations Analysis* (Englewood Cliffs, N. J., 1961).
5. *Brown Shoe Co. v. United States*, 370 U. S. 294 (1962).
6. Y. BROZEN, "Significance of Profit Data for Antitrust Policy," *Antitrust Bull.*, Spring, 1969, pp. 119-39.
7. J. M. CLARK, "Toward a Concept of Workable Competition," *A.E.R.*, June, 1940, pp. 241-56.
8. N. R. COLLINS AND L. E. PRESTON, *Concentration and Price-Cost Margins in Manufacturing Industries* (Berkeley, 1968).
9. F. E. CROXTON, D. J. COWDEN, AND S. KLEIN, *Applied General Statistics*, 3rd. ed. (Englewood Cliffs, N. J., 1967).
10. FEDERAL TRADE COMMISSION, Bureau of Economics, *Industry Classification and Concentration* (Washington, 1967).
11. M. HALL AND L. WEISS, "Firm Size and Profitability," *Rev. of Econ. and Statis.*, Aug., 1967, pp. 319-31.
12. R. G. LIPSEY AND K. LANCASTER, "The General Theory of Second Best," *Rev. of Econ. Studies*, Vol. 24, 1956-57, pp. 11-32.
13. H. M. MANN, "Seller Concentration, Barriers to Entry, and Rates of Return in Thirty Industries, 1950-1960," *Rev. of Econ. and Statis.*, Aug., 1966, pp. 296-307.
14. H. M. MANN AND J. W. MEEHAN, JR., "Concentration and Profitability: An Examination of a Recent Study," *Antitrust Bull.*, Summer, 1969, pp. 385-95.
15. J. W. MARKHAM, "Current Decisions in Antitrust," *Antitrust Bull.*, Fall, 1967, pp. 851-64.
16. E. S. MASON, *Economic Concentration and the Monopoly Problem* (Cambridge, Mass., 1957).
17. J. N. MITCHELL, Address before the Georgia Bar Association in Savannah on June 6, 1969. Reprinted in Commerce Clearing House, *Trade Regulation Reporter* (Chicago, 1969), pp. 55, 505-09.
18. O. MORGENTHAU, *On The Accuracy of Economic Observations*, 2d ed. (Princeton, 1963).
19. A. PHILLIPS, "Concentration, Scale and Technological Change in Selected Manufacturing Industries, 1899-1939," *J.P.E.*, June, 1956, pp. 179-93.
20. UNITED STATES, Dept. of Commerce, *1947 Census of Manufactures*, Vol. I, Summary Statistics (Washington, 1949).
21. UNITED STATES, Dept. of Commerce, *1954 Census of Manufactures*, Vol. I, Summary Statistics (Washington, 1957).
22. UNITED STATES, *Economic Report of the President with the Annual Report of the Council of Economic Advisers* (Washington, 1969).

23. UNITED STATES, *Presidential Task Force Report on Antitrust Policy* (Washington, 1969).
24. UNITED STATES, *Studies by the Staff of the Cabinet Committee on Price Stability* (Washington, 1969).
25. UNITED STATES, *Task Force on Productivity and Competition*, published in the *Congressional Record*, June 16, 1969, p. S6472.
26. UNITED STATES SENATE, *Concentration Ratios in Manufacturing Industry, 1958* (Washington, 1962).
27. UNITED STATES SENATE, *Concentration Ratios in Manufacturing Industry, 1963* (Washington, 1966).
28. *United States v. Philadelphia National Bank*, 374 U.S. 321 (1962).
29. E. SINGER, *Antitrust Economics: Selected Legal Cases and Economic Models* (Englewood Cliffs, N. J., 1968).

INDUSTRIAL ORGANIZATION: THE ROLE AND CONTRIBUTION OF ECONOMETRICS*

By HENRY GRABOWSKI, *Yale University*
and

DENNIS MUELLER, *Cornell University*

Industrial organization traditionally has been one of the more empirically-oriented branches of economics. Many of the hypotheses peculiar to it have been inductively derived (e.g., the Berle and Means study [1]). Its literature contains a heavy dosage of case study and descriptive statistical material. Questions of the proper measurement of various phenomena, the quality of data, and index number problems also occupy a prominent place in the literature.

While the field has been somewhat slow to take up the use of econometric techniques in its empirical work, recent years have witnessed the appearance of a large number of studies on concentration, innovation, advertising, mergers, and other industrial organization subjects that have employed multiple regression analysis and similar econometric tools. This development is encouraging and suggests that the IO economist may some day be able to derive more general hypotheses and theories than have been possible up to this point. Unfortunately trends in mergers and firm diversification [2] [5] during the post-World War II period have created (or exposed) deficiencies both in the stock of theories that can be tested and in the data available for testing them. Indeed, these trends have been so broad and strong that they threaten to vitiate future efforts to apply econometric techniques to firm and industry data and force the IO empiricist back to a case-to-case approach—with the unit of investigation being the large, diversified firm. Hence, we stand in the danger of seeing the period of infancy in the application of econometrics to industrial organization coincide with its zenith, unless we are able to develop better theories and/or come up with better data than are presently available. The remainder of this paper is devoted to a discussion of these two problems.

I. *Problems Associated with a Lack of Theory*

Industrial organization is intimately concerned with matters of economic policy, ultimately attempting to provide insights into questions such as the optimal forms of market structure on

growth and efficiency grounds and the like. While the field therefore has an empirical focus, research on these questions also requires a strong underlying theoretical structure if results applicable for policy are to be meaningfully evaluated. In its limiting extremes, traditional microeconomic theory offers a basis for formulating hypotheses on such issues. Certainly, one of the most powerful input-output generators of normative results in all economic inquiry is the microeconomist's model of perfect competition. With this model at the core and a good deal of partial equilibrium analysis available for other market structures, the empiricist can draw on a considerable body of results in formulating his speculative hypotheses.

Unfortunately, as one begins to pull at the various assumptions underlying the perfectly competitive model, one finds that it is possible to rationalize almost any kind of empirical relation. Consider the range of solutions that are normally possible under an oligopolistic market structure. Suppose, for example, one is interested in investigating how oligopoly affects a firm's allocations for activities such as advertising and R and D. It is possible on a priori theoretical grounds to have a whole range of contrasting solutions including the aggressive behavior of an extreme rivalry situation, the passive behavior of the Cournot solution, and the collusive behavior of a cooperative arrangement. Moreover, the empirical conditions under which one might expect to find any of these solutions are very nebulously formulated and tend to resist quantification. They consist of factors such as the amount of the uncertainty present, the goals of the firms, their attitude toward risk, and the historical pattern of the market's development. It is, therefore, difficult to predict in advance what might happen here solely on the basis of economic theory.

These problems have been compounded by the rise in importance of the large diversified firm. Microeconomic theory is basically a theory of the one-product firm. It has little to offer us by way of testable hypotheses of how multi-industry firms behave. Nor does it provide us with a well-accepted framework for analyzing the behavior of divisions of diversified companies. Yet the only sources of data we have come in these forms. Either we employ the census data for industries,

* The authors are indebted to the National Science Foundation for financial support under grants GS-2677 and GS-2678 which greatly facilitated completion of this paper.

which are aggregates of the figures for single product firms, the major divisions of diversified firms based in the industry, and minor divisions of firms based outside of the industry, or we resort to data on firms and accept the aggregation over the many diverse products the typical corporation now produces. Indeed, there is a basic dilemma inherent in tailoring our current theories to accommodate these existing data sources; that is, any simplifying assumption we might choose to make to allow us to use one body of data, are likely to destroy the usefulness of the other.

For example, suppose we decide to assume that the diversified firm acts as a monolithic profit maximizing institution. This would imply that each of its subdivisions would act in a manner consistent with the maximization of profit for the whole enterprise. This might lead it to buy from other divisions of the parent firm, or from outside suppliers as part of a reciprocity strategy, at prices which do not maximize the division's profit. Similarly it might undertake R and D and advertising which are only profitable because of their strong complementarity with the activities of other divisions. Whatever the merits of this assumption with regard to the diversified firm, its use is obviously inconsistent with the assumption that the units of observation maximize profit when data by industries as gathered by the Census of Manufacturing are employed.

The sheer growth in the size of corporations has produced bureaucratic structures within the firm that make traditional models of entrepreneurially led or stockholder controlled firms seem frightfully obsolete. Economics and some of its sister disciplines have responded admirably to this challenge, generating a large number of interesting behavioral hypotheses based on managerial motives, organizational theory, etc. However, this proliferation of competing hypotheses only makes the empirically-oriented economist's job that much more difficult by expanding the range of plausible alternatives from which he must choose.

Most of these "managerial" models have retained the central postulate of economic rationality embodied in the maximization assumption and have merely attempted to recast it to new situations. Others have suggested, however, that a "satisficing" and rule of thumb decision-making approach is more fruitful for explaining managerial behavior in an uncertain environment. The implications of this motivational postulate for the empiricist are to fractionalize his world view completely and make his research efforts wholly inductive in nature. For this reason, most economists have chosen not to adopt this approach.

Yet anyone who has attempted to deal empirically with dynamic behavior under uncertainty will find some common bonds with this approach. For example, the stock reaction model prevalent in investment studies is at base an error learning model with close ties to the satisficer's rule of thumb decision making. Indeed, because traditional theory is basically static in nature, mainly for reasons of analytical tractability, investigation of dynamic behavior and lag structures must proceed very much on an inductive trial and error basis.

For a variety of reasons, therefore, one seldom has much beyond a few speculative presumptions when relating existing theory to the policy-oriented issues of industrial organization. What are the consequences of this lack of a general theoretical structure in our hypothesis formulations? There would seem to be at least two basic concerns. First, the lack of theory makes it very difficult to discriminate and select among the various hypotheses relevant to a particular case. A theory serves the useful function of implying certain events and ruling out others as inconsistent with its basic assumptions. Testing the consistency or inconsistency of a theory over a variety of empirical relations, then, allows one to make informed judgments on its general acceptability. If one observes a particular set of relations without a well-defined theoretical framework, however, it is difficult to discern subtle contradictions either internally or with past results and it thus becomes difficult to reject hypotheses which are not consistent with the data. Undoubtedly, this is one reason why opposing hypotheses tend to linger on with such vigor in the industrial organization literature.

A second problem, related to the one discussed above, concerns the stability over time and across various cross-sections of any relationships observed in this manner. If, for example, one observes a particular relation between, say, a market structure variable and a measure of industrial efficiency which ultimately rests on a more complex set of unknown and unobserved relations, then the observed relation is subject to unexpected changes even if the basic structure remains invariant. Unless one has a comprehensive theoretical framework for analyzing a particular structure, the generalization of results beyond the limiting confines of one's sample is not warranted.

Nevertheless, as Professor Mason reminds us, "no one who is other than eclectic, methodologically speaking, has any business in the field of industrial organization" [4]. Decisions of public policy resting on economic considerations cannot be postponed while our theories are being refined. In many policy areas, research therefore must

necessarily go forth as best it can on current analytical structures. Research of this nature must frequently be inductive in spirit and consider as many alternatives as possible over whatever data samples are available. In reporting results, the researcher in these areas should also set forth his basic research design and in particular indicate how other approaches compare with the main findings. Too often one sees only the "tip of an iceberg" of this type of empirical analysis, preceded by an elegant but artificial attempt to make the study appear deductive in nature. However, in such circumstances, it is much more instructive to know, at least in a general way, what is the explanatory power of other variables and functional specifications. Comparison of results with more "naïve" models is also highly desirable and frequently ignored in these circumstances.

At the same time, empirical research in IO will most certainly benefit in the future as it has in the past from continued research efforts on a number of analytical frontiers. Given the significant current trends toward large size and diversification in our economy, a high priority certainly exists for theoretical studies investigating the firm in a more disaggregate fashion and ones delving into how the pieces of such organizations fit together to form rational decision-making entities. It would seem that only by gaining a greater knowledge of what is happening at the very "micro" level can we begin to make any real progress in understanding relations at the market structure level. While this kind of disaggregate approach is likely to generate theoretical models involving very complex sets of relations, simultaneous equation techniques are becoming increasingly available to help facilitate the use of these models in econometric applications. On the other hand, obtaining data at the intrafirm level to the degree required by such models is a serious problem as data sources now stand and one which will be intensively considered below.

II. *Problems Associated with the Lack of Data*

Data problems are both universally acknowledged and ignored by economists. The individual scholar is at the mercy of the private firm or government bureau for most of his data, for he possesses neither the financial nor in some cases legal power to collect them himself. Even if he is astute enough to choose the best data available for his particular model, they are frequently inadequate for the task. Rather than abandon his model to the limbo of untested hypotheses, the economist typically "adjusts" the data and truncates the model to the point where an uncomfortable compromise between theory and data is

met. The results that emerge from all of this are rightly regarded as only tentative and at best indicative of the potential for further, more intensive research on the hypothesis. Unfortunately, such research is seldom undertaken, since data of higher quality than were used in the original study rarely appear.

The industrial organization economist probably has been no more guilty of committing these sins than his colleagues in other branches of economics. However, high levels of merger activity in recent years and the trend toward conglomerate organizational entities have rapidly increased the magnitude of data problems in this area.

Mergers can effectively destroy data in both cross-section and time series studies. In time series work a merger may change the characteristics of the firm sufficiently to preclude entirely the combining of the firm's data for before and after the merger. At a minimum it forces the analyst to give up one or more degrees of freedom by deleting an observation or adding dummy variables to the model. Mergers can create similar problems when cross-section data are used. If lagged variables are employed, the merging firm must be dropped from the sample in the year of the merger. Since the number of firms in the sample and their identities will thus vary over time, the researcher's ability to draw inferences by comparing the results for homogeneous cross-sections is reduced or eliminated. Even when no lagged variables are employed, the potential for comparing cross-sections over time may be impaired by the changing composition of the sample as a result of mergers.

The increasing extent of firm diversification across product and industry lines also creates serious data problems. The implicit assumption in cross-section studies is that all of the firms in the sample follow similar behavioral patterns. This assumption seems most plausible if all of the firms produce a single well-defined product and least appealing if they each manufacture a different array of products only a few of which may be produced by the other firms with which they are grouped.

The empiricist who employs Census of Manufacturing data on industries faces problems different in kind but similar in nature to those cited above. For example, in addition to the usual aggregation problems he now must contend with the fact that his data are aggregations of whole firms, divisions of firms, and parts of divisions of firms. If cross-section regressions are estimated, he must contend with the fact that parts of a single firm may be included in a number of industries and hence that the observations are not strictly

independent. And so it goes. These are problems with which we are all familiar, and yet all ignore.

Many researchers have tried to avoid some of these problems by simply dropping the troublesome observations (firms or industries) from their sample. While this may have been a reasonable procedure as long as these phenomena occurred with some rarity, this no longer is the case. At a minimum this procedure is now very costly in terms of lost degrees of freedom. More importantly, the remaining samples of nonmerging and undiversified firms or industries with only single-product firms can no longer be assumed to be in any sense randomly generated. Indeed, they are rather peculiar groups of enterprises that seem intent upon swimming against current economic trends.

III. Some Examples

The extent to which these deficiencies in data combine to frustrate the attempts to test hypotheses is probably best illustrated by relating them to an area in which the authors have participated: the recent proliferation of studies testing the Schumpeterian hypothesis. The central question being investigated here is whether the level of innovative activity tends to increase more or less proportionately than size in a particular industry or line of industrial activity. Microeconomic theory provides conflicting strands of thought concerning the expected relationship here and, therefore, an empirical inquiry into the structural relations across various industry samples and time periods exemplifies the inductive approach suggested for such situations above.

The most desirable measure of a firm's inputs to innovational activity for such a study is data regarding the level and character of R and D expenditures, standardized by some uniform definitional criteria. The National Science Foundation has been collecting such data since 1957 for a very large sample of firms but they are able to provide information on it only aggregated over broad industrial classifications, roughly comparable to the SIC 2-digit level categories. These reported data are therefore not very useful in applications at the microeconomic level indicated above. Many firms publicly report their annual R and D expenditures, but usually say little concerning the nature of such expenses or how their definitions relate to the NSF criteria. The researcher in this area must therefore work through an array of questionable figures often surrounded by a strong sense of firm security and secrecy. Even if he is successful in obtaining R and D data privately from firms, he will have problems getting a complete or representative sample in any industry.

Indeed, for reasons of keeping the data confidential, he usually can identify the characteristics of his sample only by phrases like "ten large chemical firms" or the like.

The aggregate character of data, even at the corporate level, will provide obstacles in such a study. For example, the size variable normally available will be the firm sales or assets across all the industries in which it produces rather than the corresponding measure in its major industrial activity. Hence, unless diversification by firms into more or less research intensive industries activities is negligible, or if not, at least independent of size, this hybrid measure will influence the Schumpeterian relations estimated for each industry grouping. Where diversification is substantial, the resulting estimate may bear little relation to the true structural relation pertaining to a particular kind of industrial activity.

To help illustrate these points let us take a closer look at the firms in the chemical industry. *Moody's* for 1969 lists 183 firms which produce chemicals. Of these 183 firms, 101 can be classified as based outside of the industry, or in the case of the conglomerates, as having no real base industry at all. These range in scope from companies whose bases are in related chemical processing industries such as petroleum refining, rubber products, and drugs to companies whose base industry involves a distinctly nonchemical technology such as electronics, alarm systems, shoe machinery, and uranium mining. Obviously with over half of the firms producing chemicals being based outside of the industry, any assumptions regarding behavioral homogeneity for all chemical producing firms are quite hazardous. Even among the 82 firms that are based in the chemical industry there is a great range in the products and constellations of products they manufacture. Park Chemical, for example, is a fairly specialized manufacturer of metallurgical and automotive chemicals. Union Carbide also produces metallurgical and automotive chemicals. In addition, however, it produces a whole line of chemicals that includes agricultural pesticides, alcohols, fluorocarbon propellants and refrigerants, fumigants, latexes for paints, lubricants, plasticizers, silicones, oils, resins, urethane foams, etc. Moreover, the 1966 *Fortune Plant and Product Directory* indicates that it produces products in no less than 22 different 3-digit industries outside of chemicals, and it has divisions manufacturing carbon products, consumer products, electronics, foods, gases and equipment, metals and minerals, nuclear products, plastics, and textile materials. Despite the chemical base they share in common there is probably a greater degree of homogeneity in the product mix of

Union Carbide and some of the large diversified petroleum and acknowledged conglomerate companies than there is in the product lines of Union Carbide and Park Chemical.

Nor are things much better if one limits one's sample to the dozen or so largest firms. While each large chemical company produces any single product in common with a number of the other large chemical firms, the array and rankings of firms differs from product to product. Hence, it is difficult to speak of a single leader for the chemical industry although DuPont is by far the largest. Nor is it possible to predict which firms are most likely to introduce innovations in a specific area or what and from whom reactions can be expected.

Given these facts, it would appear very hazardous to set forth an estimated structural relation on the Schumpeterian question based on a regression over aggregate corporate figures for the largest dozen or so chemically-based firms. As shown above, such a procedure ignores the significant contributions of nonchemically based firms and includes the nonchemical production of those based in the industry.¹ Indeed, the pointed discussions and debates over the validity of observed results which have turned on questions of regression specifications and proper deflating procedures and the like (for a survey here, see [3]) may very well be academic, given the weaknesses of the data bases common to all these studies.

IV. Some Recommendations

By focusing attention on current trends in concentration, the publication of Berle and Means's classic study in 1932 [1] helped launch systematic efforts to gather data on industry and aggregate concentration. In this we are very fortunate. The important questions concerning the absolute levels and trends in industrial concentration are among the best documented issues in the field.

If similar progress is to be made on the many other hypotheses that warrant testing, then data appropriate for testing them must be collected and made available to researchers in the field. Of

particular importance here are data on firm operations on a fairly disaggregated basis. In seeking these data we have a potentially powerful ally in the stockholder. Both the stockholder and the economist have exhibited a heightened interest of late in finding out precisely what is taking place within the sprawling diversified corporation.

For this reason we would like to join in supporting the suggestion that a large data bank be established on a national level. In an age where great technological advances have been made in the data communication and processing field, one could certainly devise a system whereby such data could be made available for analysis, while at the same time respecting their confidential nature at the individual firm level. A computer system hookup from a national data bank to various academic centers which would involve central processing and the communication of results without any sight of certain confidential data inputs is an obvious development which is now feasible and highly desirable if the rational development of public policies is to be encouraged.

Barring the creation of a supply of data of this proportion, econometric work in industrial organization can be expected to move forward rather slowly. Indeed, a return in emphasis to the case study approach might even be in order—with the subject of each case study being a single large diversified firm. These organizations each possess a sufficient amount of economic importance to warrant such special treatment. And with the analytical tools and information currently available to us only a very thorough analysis of each giant firm can be expected to yield convincing conclusions. Certainly, the days when one can toss the total sales, profit, capital investment, etc., figures for G.M., 3M, Sperry Rand, and the other diversified giants into the electronic hopper and view the results with any confidence are numbered.

REFERENCES

1. A. A. BERLE AND G. C. MEANS, *The Modern Corporation and Private Property* (New York, 1932).
2. H. F. HOUGHTON, statement in *Economic Concentration*, Hearings before the Subcommittee on Antitrust and Monopoly of the Committee on the Judiciary (Washington, 1964), pp. 156-78.
3. J. W. MARKHAM, "Market Structure, Business Conduct and Innovation," *A.E.R.*, May, 1965, pp. 323-32.
4. E. S. MASON, *Economic Concentration and the Monopoly Problem* (Cambridge, Mass., 1957).
5. W. F. MUELLER, statement in *Economic Concentration*, *op. cit.*, pp. 502-18.
6. U. S. Bureau of the Census, *Concentration Ratios in Manufacturing Industry, 1963*, Part II (Washington, 1966).

¹ Some indication of the extent of intra- and inter-industry coverage to be expected in a sample of large chemically-based firms is provided from Bureau of Census concentration ratio figures for 1963 [6]. For example, the percent of the value of shipments accounted for by the eight largest firms whose primary enterprise category code is 28A (basic chemicals, plastics, synthetics) is only 36 percent in organic chemicals, 25 percent in inorganic chemicals, and 30 percent in plastics [6, p. 289]. At the same time, this source indicates that for the 16 largest firms in 28A enterprise category, approximately 30 percent of value added and 40 percent of total employment is in a different enterprise category [6, p. 310].

DISCUSSION

PETER ASCH: It should come as no surprise to anyone to find that there are some basic problems remaining in industrial organization. What is surprising is the extent to which the papers we have heard agree on the nature and causes of these problems. As industrial organization specialists we might want to consider whether there may have been a conspiracy among the authors. As economists, however, I hope we would conclude that authors, like businessmen, may independently arrive at similar conclusions, at least part of the time.

These papers argue that our basic concerns are empirical. And that our difficulties arise both from inadequate data and from a theoretical apparatus that may not permit very clear predictions to be made about firm and industry behavior. Given these kinds of problems, how are we to proceed? Here the authors part company, but only to a degree.

Dean Grether tells us that we must broaden our sights and examine more closely the internal factors in the organization and decision processes of the firm. This kind of extension, when combined with the traditional market-structure approach, may provide a theoretical basis sufficient to deal even with large and diversified companies.

Professors Grabowski and Mueller find the fact of diversification rather more troublesome. The world has changed, they state, in ways that reduce our ability to predict on the basis of orthodox theory; or, alternatively, it has changed in ways that expose theoretical inadequacies which existed previously. Despite this, they seem to believe that our theoretical problems are serious rather than fatal; and that collection of more appropriate data will enable meaningful empirical work to progress, albeit under less than ideal conditions.

Dr. Singer is at once less satisfied with our theoretical constructs and more distressed by our past empirical efforts than the other authors. If I read him correctly, he is saying that our theories of firm and industry behavior have never had much relevance; and that much empirical work, especially at the industry level, has been not merely irrelevant but misleading. The catalogue of sins that Dr. Singer finds is impressive. It includes: the application of highly restrictive "static" models to a "dynamic" world; the testing of these models in extremely simplistic formulations, using grossly inadequate data; the inappropriate employment of a partial equilibrium approach; and the use of results based on these faulty procedures to draw strong and unqualified conclusions for public policy.

I think it is important to consider what our past empirical efforts have and have not accomplished, and the issue of concentration and profit rates is a useful case in point. There is a theoretical expectation that profit rates will be higher in monopolistic industries than in competitive industries. Taking high concentration as an indicator of monopoly power, many investigators have tried to determine whether this expectation is supported empirically. The results have been mixed, but it seems safe to say that there now does exist a fairly substantial body of evidence linking profit rates and concentration.

There are of course innumerable difficulties associated with such evidence. In the first place, the theoretical implication is very imprecise. We know that a monopolist may earn an excessive profit (although this is not inevitable), whereas a competitive industry cannot. This is as much as the theory tells us, but it is not what empirical investigators have been testing. They have, by and large, addressed themselves to a similar sounding but significantly different proposition: that either the probability of an abnormal return or the size of the abnormality (or both), will increase with something called the "degree of monopoly." It may be argued with equal logic (or illogic) that this is or is not a reasonable expectation. It does not, however, follow unambiguously from the theory. And, more important, it places us in the uncomfortable position described by Grabowski and Mueller, of testing hypotheses that lack a clear theoretical foundation.

The empirical problems, both in data and in concept, that arise in studies of concentration and profit rates are often quite discouraging. Concentration indexes may be the best available measure of the "degree of monopoly" in an industry, but for many familiar reasons they are at best crude and incomplete. Indeed, it is difficult to believe that a garden variety concentration ratio could do a terribly good job of explaining profit rate variations across industries. Even if the ratio were a perfect measure of monopoly power, rates of return are strongly—perhaps primarily—influenced by other factors. These factors also may not be subject to precise measurement; yet we need to take account of them in order to observe the pure effect of concentration on profits.

The profit data commonly employed in empirical studies are also deficient. They reflect different accounting procedures and represent accounting profits rather than the economic profits with which we are really concerned. When based upon some measures of investment, profit rates may be influenced by such factors as the way in which different firms finance their operations. Finally, the observed profits attributed to any single industry may represent the aggregate of firms' activities across various industries.

When we add to all this the fact that many industries for which data are available are not theoretically well defined, the empirical problems loom large. To be sure, certain of these difficulties may be avoided by the careful investigator; but, as Professors Grabowski and Mueller point out, this may be accomplished only at some cost. An attempt to confine a study, for example, to well-defined industries and specialized firms implies decidedly nonrandom samples and makes any generalization of results all but impossible. Too often, we are probably not sufficiently vigilant about our procedures. Dr. Singer has pointed to some real problems in past studies,¹ although I cannot agree with all his objections.²

¹ For example, the use of questionable industry classifications.

² Singer shows, e.g., that the profit rate differential between high- and low-concentration industry groups

Once all these problems have been cited, what is the meaning and value of the findings on concentration and profit rates? One view, I suppose, is that the results have been based on such faulty procedures that they have no meaning whatever. I do not know whether this is what Dr. Singer believes, but if it is I must disagree with him. It is quite true that the evidence we have is far from overwhelming. The observed relationships are neither precise nor reliable across various samples and are in any case subject to heavy qualification. But it would not appear that the findings can be dismissed completely. Many of the conceptual and empirical problems mentioned above would tend to work against the observation of significant profit-concentration relationships; and investigators have generally been cognizant of difficulties that would bias conclusions in favor of their hypotheses. The empirical evidence does create some suspicion that monopolistic industries misallocate resources. This is a very modest conclusion. Indeed, the truly unfortunate fact is that the evidence does not support anything stronger than a general statement which is, on a priori grounds, quite obvious.

New problems arise when we try to interpret the findings on profit rates and concentration more closely. There are some grounds for suspecting, e.g., that concentration may be serving as a proxy for other determinants of profits. This may be one reason why past results have been inconsistent; and not nearly enough has been done to explore such an empirically susceptible possibility.

Past studies of concentration and profit rates (or industry performance generally) certainly need to be criticized. It goes without saying that careful measurement is better than sloppy measurement, and that appropriate caveats should be stated clearly along with statistical findings. To the extent that we are deficient in these areas, criticism is not only helpful but is indispensable. I believe, however, that on the whole it is the use to which past findings have been put and not the findings themselves that are the main source of difficulty.

There are two kinds of reactions that cause particular difficulty, and I am certain that everyone familiar with the industrial organization literature has witnessed both. On the one hand, there is the use of basic theory and limited evidence as the basis for an indiscriminate attack upon industry concentration. This is what bothers Dr. Singer, and I could not agree with him more. Not only does such a reaction credit theory and evidence with more than they can support, but, in the concentration-profit rate case, it ignores the obvious possibility that profits may reflect socially desirable activity.

On the other hand, there is the roughly opposite

found by Professor Mann, disappears when the individual industry observations are weighted by industry size. This is an interesting point, but I would suggest that the weighted and unweighted averages convey quite different kinds of information; and, further, that the unweighted averages employed by Mann may well be the more relevant of the two for the questions he addresses.

reaction that because we have somehow failed to prove that concentration is "bad," it is therefore "good." This reaction usually shows up, explicitly or implicitly, as an argument that inconclusive theory and evidence somehow support the *status quo*. These two reactions are equally incorrect. The unhappy fact is that our theory and evidence on the behavior of firms and industries are inconclusive. Neither a decision to restructure the economy nor a decision not to restructure it can be predicated on the knowledge we have gathered to date.

I find this a disturbing conclusion. After all, it has been thirty-two years since Edward S. Mason noted that the effects of economics upon public competition policy are largely negative. Our empirical techniques have become more sophisticated, our data sources have improved somewhat, and we have made very extensive use of computers. But I would find it difficult today to discuss our impact on policy in much more optimistic terms than did Mason.

I have concentrated my remarks on some of the points raised by the Singer paper, but I would like to mention one issue concerning all three papers. I have one objection, which is perhaps something of a quibble. This is the impression, if not the implication, which might be drawn from the papers as a group, that better measurement will somehow enable us to solve all relevant problems in industrial organization. None of the authors says this, and they may not believe it; but it is possible to conclude from what has been said and from what has not been said that we have no other pressing concerns.

It should be pointed out that even with appropriate theories and ideal data some problems in which we are all interested would remain. This can be seen most generally in discussion of public policy issues. I am most unhappily impressed by the extent to which our policymakers are confused about priorities, both in antitrust and the regulated sectors. It is not just a matter of lacking good economic information—although this is an important part of the problem—but of failing to see that noble objectives inevitably conflict and that there is a clear need to decide in effect what is important to us. Economics cannot define the "right" objectives; but economists can contribute to the discussion. Unless we do so, even our best theoretical and empirical efforts ultimately may count for very little.

ALMARIN PHILLIPS: The evaluations of the state of industrial organization in the papers by E. T. Grether, Eugene M. Singer, and Henry Grabowski and Dennis Mueller are much more alike than their titles suggest.

In each paper, occasion is found to note the inadequacies of the theoretical foundations on which industrial organization studies are based. Dean Grether resists a direct suggestion that the market structure-conduct-performance approach to these studies be abandoned. Yet he attacks the approach with some vigor. In one sweep, he (a) acknowledges that the approach is based both logically and historically on the revolution in theory which took place in the 1930's; (b) argues that the approach is largely irrelevant for studies of the "most important issue for the field" (i.e., the behavior of the large diversified corporation); (c) holds

that the "general [market] case . . . really is oligopoly"; and (d) laments "the inherent negative stance of most oligopoly models"—on which the structure-conduct-performance approach is admittedly based.

Dr. Singer finds little help in industrial organization studies from either theories of competition or monopoly. In his words, "a model can suggest almost any alternative." But, in Singer's view, whether helpful or not, the ostensible use of models in making predictions or policy decisions concerning market performance by picking, say, a concentration ratio as the measure of structure and, say, profits as the measure of performance is a backwards step for the field. Further, even if the specific predictions were accurate, problems such as those appearing in the theory of second-best would cloud their welfare implications.

Professors Grabowski and Mueller come directly to the criticism of theory; they devote the first substantive section of their paper to "Problems Associated with a Lack of Theory." Among other things, they join Dean Grether in arguing the importance of the large diversified firm in today's world and in asserting the irrelevance of nearly all existing microtheory with respect to the behavior of such firms. Considering all three papers, one gains a sense of considerable despair.

The papers also note problems of an empirical nature. Dean Grether observes that Census data "are relatively crude" and fail to cover nearly the gamut of the varied structural and performance dimensions. Dr. Singer reviews the evidence on trends in concentration, pondering the economic import of the findings and demonstrating their sensitivity to alternative weighting schemes. Similar sensitivity is shown for measures of relationships between concentration and profit rates. Professors Grabowski and Mueller devote the second substantive portion of their paper to "Problems Associated with the Lack of Data." Again, the sense of despair.

There is little in the specifics of these papers with which I am inclined to disagree. Oh, there are matters of emphasis, interpretations of fact, omissions of certain references I might have used, but these are hardly things to quibble over here. It is the pessimistic tone set by the papers—especially the Singer and Grabowski and Mueller papers—that is troubling. While it is certainly true that theoretical and data deficiencies impair the economics of industrial organization, my view is that the picture is not so bleak as might be supposed from the present interpretations of the state of the field. Interesting operational hypotheses remain within the present state and new directions for the field seem reasonably clear.

Dean Grether, the most sanguine of the present authors, indicates what some of the changes might be. He says (by implication, at least) that the structure-conduct-performance approach requires modification and amendment. Research is needed on the internal organization and behavior of large firms, since structure fails to explain their behavior. Among the subjects of inquiry would be "the investigation of the interaction of structural and behavioral factors into long-period analysis so that we would have a better understanding of the long-run determination of market structures."

Thus, as I read him, Dean Grether would combine the static structure-conduct-performance approach with dynamic elements that treat structure and conduct as being functionally related over time to differences in performance. It seems that Grabowski and Mueller adopt something of the same position in their discussion of dynamic behavior and lag structures, though this is not wholly clear. In any case, I would heartily endorse research efforts in such a direction.

The Grether paper (not strongly, yet by more than implication) also urges additional research based on organizational theories and behavioral and management sciences. Grabowski and Mueller again seem to sympathize. None of the papers, however, stresses the complementarity between research on dynamic variations of the structure-conduct-performance approach and research on large-scale organizations, group behavior and group decision making. The latter is essential to the erection of dynamic models incorporating adaptive responses.

Some studies following these lines have already been made. There is scope for many more. Indeed, in the face of the propelling policy interests and applications which Grether sees, additional and fruitful research is all but inevitable. For the next few years, at least, research relating to policy aspects of microeconomics will gain in importance. If industrial organization is to make its contribution, however, it will have to broaden its areas of topical interests as well as strengthen its methods. Antitrust and the regulation of business through controls on structure and conduct will not, in my opinion, be the vital concerns. Instead, improvements in resource allocation will come more from sagacious intervention in the market process through myriad forms of other controls, including public pricing and tax-subsidy techniques that account for externalities. Here industrial organization and the microeconomics of public finance could well become indistinguishable. And economists could make major contributions to the commonweal.

JAMES W. MCKIE: The observations in these three papers are diffuse, no doubt reflecting the diffuseness of industrial organization itself. Without attempting to discuss all of the cardinal points made by their authors, I would like to offer my reactions to two of them.

1. *Structure, Behavior, and Performance.* Students of industrial organization are certainly eclectic in their choice of applied theoretical models, but the choice depends largely on what questions they are attempting to answer. E. T. Grether points out that analysis often runs directly from structure to prediction of performance results, and that for this purpose only a simple specification of "structure" is required: concentration and size, entry conditions, and (sometimes) product differentiation. Some investigators have found concentration alone to be enough for prediction if performance is measured by a profits test, since the other two elements of structure are largely collinear with concentration in their effects upon profits. Other students have added tests for other structural variables, but their yield in fuller explanation of profits performance shows rapidly diminishing returns.

The leanness of structural specifications for the purpose of testing performance is simply the counterpart of the austere assumptions of theoretical market models. These models generally predict equilibrium results, of which profit performance is a key dimension. They are not concerned with describing or predicting behavior. In following their lead, students who examine the relationship between concentration (and perhaps entry conditions) and profit performance are not trying to supersede the theoretical models of monopoly and competitive equilibrium. They are trying to identify and predict the cases in the real world whose performance lies toward the one pole or the other. This analysis is ultimately based on theory—that is to say, on a general systematic view of economic relationships; it is not a manifestation of rootless, *ad hoc* empiricism. Nor is the policy implied by its findings.

Let me digress a moment in order to comment on Eugene Singer's reference to the "second-best" problem. He observes that "deconcentration of some but not all industries may merely benefit a few industries at the expense of other industries with either a negative or no net benefit conferred on the economy." True, it may; but it may not. The Antitrust Task Force last year thought that it probably would not. It also thought that the burden of proof should be shared by those who would raise this objection, though anyone can agree that we would like to know more about it. What bothers me about the second-best issue is that it seems to come up of late in practically every discussion of policy. One can scarcely propose a policy touching any part of the micro-economic structure without the red flag of "second-best" going down. If we had to correct every misallocation in the economy before proceeding against any problem, we could never have any policy for improvement at all. I suspect that the second-best principle is beginning to assume the role that the New Welfare Economics used to have: that of paralyzing the will to attempt solutions to outstanding economic problems. The point is that there are third and fourth bests, ranging down to the worst, or perhaps the second worst. We want to make sure, of course, that offsetting changes elsewhere do not rob us of the ostensible benefits of any policy change, but we should not assume that any problem we are examining actually represents a second-best solution.

Besides predicting performance, economic analysis should be able to explain the market behavior that we actually observe. For this purpose we need a more complex and detailed specification of structure than the specification that suffices for performance analysis. Behavioral problems are interesting chiefly in disequilibrium systems, or when our general analysis fails to predict a market equilibrium. This is eminently true of oligopoly, and it is no accident that industrial organiza-

tion has been preoccupied with oligopolistic markets. "Pure" oligopoly theory itself is highly eclectic and evidently conscious of its own inadequacy. There are a great many patterns to explain and many combinations of predictive elements are possible.

2. *Management Motivation and the Behavior of the Large Corporation.* Dissatisfaction with profit maximizing as an all-embracing principle of business behavior is not new, nor did it grow out of the study of industrial organization. Economists, like everyone else, have known for a long time that managers' interests tend to deviate somewhat from stockholders' interests in the large corporation, and that these organizations suffer to some degree from suboptimizing by their component parts. We are also aware of the difference between profit maximizing in the short run and in the long run and of the difficulty of defining maximizing at all in the presence of uncertainty and differing risk preferences. These are at least frictions, helping to account for observed deviations from equilibrium. At best they lead to alternative hypotheses, such as "sales maximization subject to a minimum profit constraint."

Yet the profit motive does not seem to have been replaced altogether by these elaborations. I agree with the observation of Grabowski and Mueller that an uncritical acceptance of "satisficing" as a postulate of motivation would "fractionalize" the economist's world completely and reduce him to marveling at the drama of business executives each doing his thing. But it is striking how often the alternative motives seem to evaporate in actual cases—how satisfied the executives are with profits and how dissatisfied without. We can say, I think, that there is a lot more leeway for secondary motivation in noncompetitive situations than when the pressures of competition force businessmen to minimize costs and struggle for their share of capital resources. This has often been observed.

The large conglomerate firm presents a more complex empirical problem for explaining behavior than the large single-industry corporation but not one different in kind. The policy implications also seem familiar, if economic performance is defined in the usual manner: to keep enough competitive pressure on conglomerates and their component parts to ensure that the survival motive, the profit motive, the sales-maximizing motive, and the satisficing motive all lead the firm to minimize costs, to innovate, to maximize efficiency, and to direct the flow of capital resources into the activities most urgently demanded by the economy. If the growth of conglomerates is artificially encouraged by aberrations and anomalies in capital markets, we should be able to find the tools of policy to deal with that problem, also. New departures in antitrust policy may well be needed, and certainly we need better information about conglomerates than we have now.

NEW DIMENSIONS IN NATURAL RESOURCES

NEW DIMENSIONS FOR TRANSNATIONAL MARINE RESOURCES

By FRANCIS T. CHRISTY, JR.
Resources for the Future, Inc.

Introduction

It is not difficult to anticipate that the oceans will become a source of major conflict between nations during the next decade. The conditions that have prevented states from struggling over the sea's wealth have changed drastically in the past few years. In part, this is due to the growing value of marine resources and the increasing incentive of nations to appropriate the resources. And in part it is due to the greatly decreasing international restraints to the extension of national limits and other means for the unilateral acquisition of wealth.

In 1945, the Truman Proclamation extended U.S. rights to seabed resources out to the edge of the continental shelf. Most states followed suit. And now rights are being extended directly by claims and indirectly by the granting of concessions beyond the edge of the shelf and down the continental slope.

In 1958, when Iceland claimed exclusive fishing limits out to twelve miles, it was over the opposition of many maritime states. But the success of the claim and the acceptance by the British broke the dam of the traditional narrow limits, and most states were quick to adopt similar extensions. Several states have actually gone well beyond twelve miles, and others may soon follow.

In these and other less direct ways the traditional restraints to national appropriation of the sea's wealth are being broken. The continued viability of the principle of the freedom of the seas is open to question. And as the seas' wealth becomes divided, both along its margins and in the distant regions, conflict becomes inevitable.

The alleviation of conflict is possible, but it will require, among other things, some viable substitute for the freedom of the seas so that there will be new restraints to—or at least an orderly process for—the appropriation of exclusive use rights. And this, in turn, will require the formulation of new international rules and institutions.

Social scientists have an important role to play in the formulation of the new rules and institutions. But thus far, only a handful of social scientists have done significant research on the problems of international marine resources and policies. This paper is written in the hope that it

may tempt more social scientists to turn to the oceans.

The central differences between resources of land and those of the sea are due to the different political and natural media in which the resources occur: (a) the absence of national jurisdiction and (b) the overlap of uses within the same physical space. Five elements help to characterize the differences. First, exclusive rights cannot be readily obtained and it is extremely difficult to control the amounts of capital and labor applied to the resource. A second element lies in the direct confrontation in the exploitation of the same resource between economic units that operate under different wage/price structures and different value systems. The third, already mentioned above, is the relaxation of restraints to the extension and appropriation of rights of ownership. The fourth is the difficulty of management. And the fifth, which is a product of the natural medium, lies in the fact that the resources occur in three dimensional space and are strongly influenced by the flow of forces and currents through space.

Open Access

Resources for which satisfactory exclusive rights are not available or are difficult to obtain are generally referred to as open access or common property resources. Under certain types of management, these may include air, water, outdoor recreation areas, oil pools, fish, wildlife, scenic areas, etc. In some cases, the use of the resource is incidental to the production of other goods as, for example, when water is used as cooling agent in the production of steel or air is used for the disposal of gaseous wastes. In these situations, economists have tended to analyze the common property situations in terms of externalities.

In other cases, however, the resource itself is the major factor of production: common pools of oil, fish, wildlife, etc. Here the research approach is on the conditions and consequences of open access rather than on externalities. While there is no sharp distinction between land and sea resources in this respect, the condition of open access is more prevalent and more difficult to change on the oceans than on land. This is in part due to the

fugitive characteristics of some of the resources—the fact that salmon, tuna, and even the plant nutrients may travel thousands of miles. But it is also due to the traditional difficulty of closing access to a resource in which all nations feel that they have a right.

Since the exposition in 1954 by H. Scott Gordon [1] of the theory of common property in fisheries, there have been few studies of either an empirical or theoretical nature. The few fisheries that have been studied, by Crutchfield and others [2] [3] [4] [5] [6], have clearly demonstrated the prevalence of economic waste associated with open access. But the studies are both few in number and limited in significance. They have dealt only with single fisheries and the efforts of single nations. Research should be expanded to include studies of the economic consequences of open access to fisheries that are ecologically interrelated, of fishing industries that utilize different species during different seasons, and of fisheries that are utilized by nations with different wage/price structures. Such studies will require sophisticated handling of poor data as well as the ability to deal with such peculiarities of fisheries as the share system of wages, severe fluctuations in annual yields, the relatively high degree of immobility of many fishermen, and the strange but common forms of conservation regulations that prohibit technological innovations and increase catching costs.

In addition to empirical analyses, there is considerable need for expansion and refinement of the theoretical elements of common property fisheries.¹ Models simulating the conditions described above could be particularly useful in pointing the way towards improved management under conditions of open access. Short-term effects, such as intraseasonally declining yield functions and changes in effort, should also be analyzed.

In addition to fisheries, the principle of the freedom of the seas impedes the acquisition of exclusive rights for other uses and resources of the sea. Several large companies, such as Tenneco, International Nickel Company, and Kennecott Copper, have given evidence of a strong interest in the possibility of exploiting the manganese nodules that lie on the deep ocean floor, well beyond the limits of present coastal state jurisdic-

tion. While the likelihood of commercial success still appears to be a decade or two away [9] questions are being raised as to the desirability of exclusive rights as against open access. The significance of the questions lies in the fact that if exclusive rights are desirable, it will necessitate either some totally new international machinery or a division of the seabed among coastal states.

Disparity of National Values

A second, and sharper, distinction between land and sea resources lies in the direct confrontation between units that operate under highly diverse national values and economies. On the Grand Banks of the Northwest Atlantic, for example, the fishermen of more than a dozen nations congregate to catch cod, haddock, and other groundfish. They come in everything from small dories to highly organized fleets of specialized vessels, from the local "outports" of Newfoundland and the distant cities of Japan.

The freedom of the seas has permitted a vast distribution of fishing effort. The Japanese and Soviets fish in all corners of the world with large factory vessels and special catching and transport ships. The South Koreans have the largest tuna fleet in the South Atlantic. Even the U.S., as undeveloped a fishing nation as it is, has vessels off Latin America fishing for tuna and shrimp.

And since the major fishing grounds of the world tend to be limited in size and number and generally close to land, the distant water fishermen are frequently in direct competition with the small local craft of the nearby coastal states. The difficulties occur when, as is inevitable, the stocks become depleted and regulation becomes necessary. At this point, it becomes important to determine the value of the stocks and the effect of alternative regulations upon the economic interests of the users—a task that is far from easy in view of the disparity of the interests.

On a much broader scale, differing national values and interests will have to be sorted out as the world community attempts to give meaning to such elusive terms as the "common heritage" or "common wealth" of the seas. Two years ago, the General Assembly of the United Nations created a special seabed committee in response to the growing deep water activities of oil companies, the possible exploitation of deep sea manganese nodules, and the possible use of the seabed for the placement of nuclear weapons. The two economic questions that are now being raised deal, first, with the appropriate and acceptable limits of coastal state jurisdiction and, second, with the character of the international regime that should

¹ For a general discussion of research needs see Clifford S. Russell [7]. For recent theoretical analyses see Vernon Smith [8], Harlan C. Lampe in [4], and the papers presented at the MacMillan Fisheries Economics Symposium held by A. Scott, Univ. of British Columbia, Mar. 24–25, 1969.

govern exploitation beyond those limits. Although these questions are directly concerned with the production and distribution of the sea's wealth, little has been done to determine the nature of that wealth, much less the effects of different limits and regimes upon its distribution or upon the concept of the "common heritage."

One indication of the importance of such studies lies in the fact that those states that oppose international machinery do so in large part because they feel that "the concept of a 'common heritage' [is] unrealistic" since it disregards the "existence of States with differing social systems and differing systems for the ownership of property" [10].

Distribution of Wealth

A third difference between land and sea resources lies in the fact that there is no system for the orderly acquisition of the sea's wealth. Through various means, direct and indirect, states are acquiring greater shares of the sea's wealth for themselves and reducing the amount of wealth available to others. Since the reductions have, thus far, been relatively small or widely diffused, the conflicts have also been relatively insignificant. But as the incentives to acquire greater shares of the sea's wealth increase, so will the severity of the conflicts.

As already stated, there are no clear-cut or universally accepted limits to the jurisdiction of coastal states. It has been suggested that no limits should be drawn short of the midpoints of the oceans so that all areas of the sea will come under some nation's jurisdiction. But there are several reasons why such a solution would be both impractical and unacceptable. For one thing, fish are no respecter of boundaries, and fisheries' management would still require some kind of international agreement. For another, the strategic position of certain islands would make such a distribution a crazy patchwork of politics (viz., the South Sea Islands, Bermuda, the Bahamas, Tristan da Cunha). The U.S. Navy would oppose the solution because it would tend to impede mobility of warships. And the Soviets would not agree because they would gain so little.

In recognition of these effects, most states now agree that limits will have to be drawn short of the midpoints and that a large area of the seas should remain international in character. But there is no agreement as to where the limits should be drawn and little knowledge as to how different limits would affect the distribution of wealth.

Maintaining a portion of the seas as international, however, does not prevent states from

finding other means than the assertion of claims for the acquisition of the sea's wealth. Several stocks of marine fish and mammals have been divided up among a few parties: Antarctic whales, Fraser River salmon, North Pacific fur seals, both eastern and western Pacific salmon and king crabs, and others. In each of these cases, the division of wealth is based on the assumption that nonparticipants will continue to refrain from participating—an assumption of questionable merit.

There are also other, less obvious distributive effects taking place. Certain conservation rules, while ostensibly preserving open access, have differential effects upon the participants and operate as a barrier to new entry. In some cases, such as Antarctic whales, it has been advantageous for a few parties to decimate, if not extinguish, a resource—to the diminution of the "common heritage." The disposal of wastes (including poison gases) may be a highly valuable benefit to one state but a cost to all others. Even scientific research may have distributive effects, as pointed out by the Maltese ambassador to the United Nations. "Who," he asked, "is likely to profit most in practice, that is economically and militarily, from the scientific programmes proposed to us? Not land-locked countries, not countries bordering on closed seas, not the developing world in general—and, I would add, not the goal of demilitarization of the ocean floor" [11].

Research on the wealth of the seas and its distribution should be valuable in reducing the potential for conflict. It would be helpful to have studies that could improve perceptions of the sea's wealth, evaluate the gains and losses of alternative actions and arrangements, analyze the mechanisms by which wealth is being distributed, and, in general, provide the participants with more common and accurate bases for decisions.

International Management

For reasons that are obvious in the foregoing discussion, the management of the sea's resources is much more difficult than that for land. At the same time, however, it is also clear that the need for management systems is becoming critical, both for the purposes of facilitating rational use and for the alleviation of conflict.

Thus far, the experience on the seas has been far from satisfactory. There is only one international conservation agreement that has been economically effective. This agreement provides that the fur seals of the North Pacific can only be harvested on their breeding islands. The two states that refrain from taking seals on the high seas (Japan and Canada) receive a portion of the

skins that are harvested (by the U.S. and U.S.S.R.) in exchange for their loss of access. In all other international agreements, the regulations are designed only to prevent physical waste. Of these, only a few have even succeeded in reaching this objective, generally at great cost.

For example, the Inter-American Tropical Tuna Commission recently agreed to impose a total limit on the seasonal catch of yellowfin tuna. As might have been expected from the experience of a similar control established on Pacific halibut in the 1930's, the length of season has dropped rapidly as the fishermen have increased their investments in larger and faster vessels, in attempts to get greater shares for themselves before the total limit is reached and the season closes. When the limit was imposed in 1967, the capacity of the U.S. tropical tuna fleet was 37,000 tons. By the start of the 1968 season, it was up to 43,600 tons. At the end of that season, it was at 48,300. It is now 55,600 tons and will reach 65,000 tons in 1970—an increase of 75 percent in the four years of control. The allowable catch was increased this year by a third, over the objections of the scientists, but still the season has dropped to four months from the nine that it was formerly.

In addition to the several agreements already mentioned, there are many more to which the U.S. is a party. In the last few years, several of these have been *ad hoc* bilateral arrangements with the Soviets and the Japanese, by which they abstain from fishing in certain international waters of interest to U.S. fishermen and receive, in exchange, certain rights within the area of exclusive U.S. limits. A complex melange of overlapping and intermixed agreements is developing throughout the world. And yet there have been virtually no studies of decision making and operations of international fishery arrangements—no analyses of the mechanisms of trade, of gains and losses, constraints, objectives, or degrees of success.

Overlapping Uses

The final distinction that may be helpful in differentiating between land and sea resources derives more from the natural than the political medium. The buoyancy of water supports surface activities. The water's low density facilitates sub-surface activities. And its fluidity creates an open environmental system. In vertical dimensions, coal is mined in shafts below the sea floor, oil is drilled from platforms fixed to the bottom and rising above the water, minerals can be dredged from the surface of the ocean bed on which solid wastes (some of them toxic) may be deposited,

sedentary animals are scraped from the bed on which telephone cables may lie, bottom feeding animals are caught in traps or trawls, mid-water species may be taken by hook and line or by trawls which occasionally interfere with submarines, surface species are taken by net and harpoon, and the surface itself is used for shipping as well as the vessels engaged in extracting resources. In horizontal dimensions, the flow of water brings pollutants, plant nutrients, temperature changes, and forces from far distant areas, while fish, ships, and submarines move with or against the currents.

It might appear that the oceans are so vast that interference between these various uses and forces is not likely to be great. But for several reasons many of the uses tend to be concentrated in certain areas. On fertile fishing grounds, for example, there are conflicts between the use of stationary gear (such as traps), floating gear (such as longlines), and towed gear such as trawls. Greenland fishermen have complained that their longlines and gillnets have been destroyed by German trawlers. And the Japanese have claimed that some of their longline gear has been deliberately destroyed by Mexican purse seiners.

Fishing is generally good in the relatively shallow waters of the continental shelf. But it is also economical to drill for oil and gas in these areas, and their proximity to land means that the same areas may be heavily used by shipping. This is evident in the Gulf of Mexico and the North Sea and is developing on the Grand Banks. The problems of conflicting uses in these areas are compounded by the differences in jurisdiction—the coastal state may have control of shipping out to three miles, control of fishing out to twelve miles, and control of oil and gas operations well beyond that. Thus, there are difficulties not only in choosing among conflicting uses but also in determining the appropriate institutions and vehicles for making the choices.

Summary

Essentially, the seas are in a stage of transition. They are moving from a condition in which property rights are almost nonexistent to a condition in which property rights of some form will become appropriated or made available. The transition stage may be of short duration for certain resources along the sea's margins—or it may be much longer for the relatively abundant resources far from shore. The rights may go to individual states or to some agent acting on behalf of regional or world communities. But eventually, as

the sea's resources become more valuable, exclusive rights will be acquired.

The acquisition of exclusive rights clearly affects the distribution of the sea's wealth. And while it is conceivable that such distribution can be achieved without conflict, it is most unlikely that nations will be able to develop the inordinate amounts of good will, mutual understanding, and institutional innovations that will be necessary to do this. A more reasonable goal is to reduce the potential for conflict and minimize the damages from conflict. Social scientists can help in reaching this goal. But to do so, they will have to turn to the seas in earnest and deal with the new dimensions that mark the use of the sea's resources.

REFERENCES

1. H. SCOTT GORDON, "The Economic Theory of a Common Property Resource," *J.P.E.*, Apr., 1954, pp. 124-42.
2. JAMES A. CRUTCHFIELD AND ARNOLD ZELLNER, *Economic Aspects of the Pacific Halibut Fishery*, Fishery Ind. Res., Vol. 1 (U.S. Dept. of the Interior, 1961).
3. JAMES A. CRUTCHFIELD AND GIULIO PONTECORVO, *The Pacific Salmon Fisheries; A Study of Irrational Conservation* (Johns Hopkins Press for Resources for the Future, Inc., 1969).
4. F. W. BELL AND J. E. HAZLETON, eds., *Recent Developments and Research in Fisheries Economics* (Dobbs Ferry, New York: Oceana Publications for the New England Economic Research Foundation, 1967).
5. R. HAMLISCH, ed., *The Economic Effects of Fishery Regulation*, FAO Fishery Reports No. 5, papers presented at an expert meeting, Ottawa, 1961 (Rome, FAO, 1962).
6. R. TURVEY AND J. WISEMAN, eds., *The Economics of Fisheries* (Rome, FAO, 1957).
7. C. S. RUSSELL, "Applications of Mathematical Economics in Marine Resources Research," in L. Alexander, ed., *The Law of the Sea*, Proceedings of the Fourth Annual Conference (Kingston, R. I.: Law of the Sea Institute, forthcoming).
8. V. SMITH, "On Models of Commercial Fishing," *J.P.E.*, Jan.-Feb., 1969.
9. P. SORENSON AND W. MEAD, "A Cost-Benefit Analysis of Ocean Mineral Resources Development: The Case of Manganese Nodules," *American J. of Agric. Econ.*, Dec., 1968, pp. 1611-20.
10. Report of the Secretary-General, "Study on the Question of Establishing in Due Time Appropriate Machinery for the Promotion of the Exploration and Exploitation of the Resources of the Sea-Bed and the Ocean Floor Beyond the Limits of National Jurisdiction, and the Use of These Resources in the Interests of Mankind," United Nations General Assembly, A/AC.138/12, June 18, 1969, p. 11.
11. 23 GAOR, Provisional Verbatim Record of the 1589th Meeting of the First Committee 13-16 (U.N. Doc. No. A/C.1/PV.1589) (1968).

THE ECONOMICS OF CONGESTION AND POLLUTION: AN INTEGRATED VIEW

By JEROME ROTHENBERG
Massachusetts Institute of Technology

I. *Generic Congestion, Customary Congestion, and Pollution*

Increasing popular and professional attention is being given to two types of phenomena characterized by substantial externalities: urban congestion and environmental pollution. Both represent the unkind rub of human activities on one another, where there is no intermediation of a market to enable affected parties to confront their tormentors. Examination of the phenomena reveals that they have much in common beyond the sheer fact of externalities. The structure of the externalities is similar,¹ and an exploitation of this similarity throws light on why they are both becoming critical problems in our society, what their consequences are, and what kinds of measures can be used to mitigate their damage.

In the present paper we propose to treat congestion and pollution problems as members of a single class, integrating the distinctive characteristics of the two together as aspects of one phenomenon. We shall first discuss the nature of this more abstract configuration of socioeconomic interaction, and then display its ramifications in a mathematical model which treats both positive and normative considerations.

Highway traffic jams, queues, crowding of beaches and parks and museums, and air and water and noise pollution are all forms of social congestion. Common to all of them is that more than one agent is attempting to share a type of service that is not furnished in a separable unit earmarked for each user. They are consuming in common form a "public good"—whether it be a highway or beach or volume of air or a water-course. In all, the presence of other users adversely affects the quality of services which that public good renders to each. Quality deterioration may be revealed in terms of the length of time, safety or psychological tension of an auto trip, the size of available area of picnic or swim space, the level of aesthetic disfigurement of a setting by

litter or noise, the degree of obstructedness or average viewing time, of paintings, the amount of eye and lung or ear irritation in the air, or the odor, taste, bacterial count or fish population in a body of water. In all of these, moreover, quality deterioration does not set in as soon as more than one user attempts to share the public good in question. Depending on the nature of the public good, a differing but rather wide range of users may be accommodated with no perceivable deterioration of quality. Each good has a "capacity," or threshold, beyond which interference effects first become noticeable and then increase disproportionately. Finally, while some part of the capacity of the public good may be natural, or given by nature as a "free good" (this does not apply to museums), deliberate human action can either increase that capacity or mitigate the quality impairment stemming from any given level of socioeconomic interaction. Thus, low flow augmentation procedures can enhance the waste assimilative powers—the capacity—of a river, or sewerage treatment plants can decrease the extent to which any given flow of effluents overwhelms it. Various forms of recycling affect air pollution, widening a highway affects traffic flow, rationing devices influence the obstructiveness of a given crowd in a swimming pool or museum.

This composite of public good sharing and the policy variability of medium capacity and rate of interference by-product for each unit flow of interaction is the basic characteristic of the generic "congestion" process which we shall treat. Under this broad head the more conventionally conceived congestion and pollution take somewhat specialized forms.

What distinguishes congestion from pollution in terms of the more inclusive concept is chiefly the relationship among the generators and victims of "interference." If highway traffic is the classic example of congestion, then the central interpersonal distributive fact about it is that all users are using the medium (the public good) in much the same way, each is damaging service quality for both others and himself, and the ratio of self to other damages is approximately the same for all users. It would be difficult to separate users into abusers and victims. Congestion is not looked on

¹ The illuminating works of Allen V. Kneese on environmental pollution—as for example his, "Air Pollution—General Background and Some Economic Prospects," chapter in Harold Wolozin, ed., *The Economics of Air Pollution* (Norton, 1966), p. 34—recognize this similarity.

as a process in which important real income redistribution occurs—some benefiting by imposing damages upon others. The whole user group loses homogeneously by its self-imposed interaction.

The essence of pollution, on the other hand, is that there are some users who do abuse the medium—the polluters—while others are relatively passive victims of such abuse—the “public.” In these processes users differ among themselves in how they use the medium. Some users employ rivers as sewers for noxious materials; others—downstream—simply want to drink the water. Some use the atmosphere similarly as a medium for noxious waste disposal; others simply want to breathe it. Jet planes make the noise; housewives are forced to submit to it. Thus, pollution often lends itself to a distinction between destructive and constructive uses of a medium, between guilty and innocent parties. Significant—whether morally or in scope—income redistribution is a key aspect of the process.

The legal and ethical characterizations of the distinction are controversial. Is a pulp processor destructive because of what he puts into the river, or because of what other upstream users are putting in at the same time, or because of what a downstream user is taking out? If no other upstream discharger existed or if there were no downstream user, his actions would hurt no one. Who has property rights in the river? Is it a question of who came first? Even if it were, this would not resolve the issue, since it is not the sheer existence of upstream and downstream users that causes the difficulty but the scale of their interaction. Up to a certain scale, upstream use causes no quality deterioration. It is the marginal increment which crosses the threshold that becomes noticed. But then the damage is produced by all the upstream users, not just the marginal one. Can a destructive-nondestructive dichotomization of uses rest on scale considerations?

Economists like Coase and Buchanan have leaned on this kind of consideration to argue an essential legal symmetry for different uses. They accept any *status quo* pollution and suggest that it is ethically and allocationally equivalent whether victims offer payments to polluters to desist or the state forces polluters to pay compensation to victims for the right to continue. Only income distribution is affected.

One can recognize that real pollution depends on primacy, on scale of polluter use, and on the presence and scale of victim use, without wishing to abandon the distinction between destructive and nondestructive uses. At any time, given the

contemporary context of the nature and volume of potential users of a medium, some uses are at least potentially damaging to other conceivable users. There is an important asymmetry between those who spew gases into the air and those who only want to breathe it. The former do at least potential ill to the latter, but the latter do not do damage to the former. If this kind of asymmetry be granted, then it is not the case that neutrality (symmetry) of property rights is allocationally neutral. For if external diseconomies against others can be expected to lead to bribes by victims to desist, then the production of negative externalities becomes a valid by-product of primary production. Profitability is enhanced whenever any firm can select from among its input and/or output alternatives those which cause substantial damage to third parties. Resource use will tend to become specialized toward much-augmented third-party interference. The new legal industry of selling protection against disturbance will be highly profitable. So long as the distinction between negative and positive externalities is maintained, such an adaptation of resource use must be deplored.

Our treatment of pollution does maintain the distinction. There is a unique direction of service quality impairment, and some uses involve a higher rate of “potential impairment” than others. Whether potential impairment eventuates as actual impairment depends on concrete situations—the scale considerations.

“Generic congestion” subsumes both customary congestion and pollution as special cases of a general phenomenon: (1) pure congestion is the case where all users generate identical rates of quality interference per unit of activity and share equally in the resulting quality impairment; (2) pure pollution is the case where some users generate very high rates of per unit interference while others generate zero rates and only the latter experience quality impairment; (3) the general case is where all users both generate impairment and share in it, but they differ from one another in both respects. The variety of both abuse and victimization prevents an easy or complete categorization of users into guilty and innocent. Generic congestion, by enabling us to study a multitude of patterns of relative generation and sharing of damages, can throw light on a greater range of allocational, distributional, and public policy issues than the customary congestion and pollution concepts. Indeed, conventional terminology makes it difficult to understand most observable real-world phenomena, because it

obtrudes extreme case insights into situations that typically have elements of both extremes. Real-world cases are neither pure congestion nor pure pollution. Policy recommendations based on these polar concepts are likely to be deficient to the extent that they fail to take account of the mixed characteristics.

In the remainder of the paper we shall briefly present a model of generic congestion to give some idea of the facets which can be explored.

II. A Model of Generic Congestion

A. *Preliminaries.* In order to place the model in more than a partial equilibrium context we assume that there are only two commodities in the system: X and Z . X is subject to generic congestion but Z is not. The nature of the generic congestion is as follows. Government provides a public good-type capacity from which it sells shared services to the population. A variety of forms of crowding and interactive interference characterize the actual consumption of these services. The population contains three different types of users. All three share equally the results of congestion, but they contribute to it unequally. There is a low polluter, a middle polluter, and a high polluter group. Each generates a low, middle, or high rate of interference for every unit of X he consumes: groups L_1 low, L_2 middle, L_3 high; $L^i \in L_j$ member i of group j ; the total population is divided among the groups in the fraction n_1, n_2, n_3 , respectively.

While we have spoken about the consequences of congestion as quality impairment, we shall facilitate the treatment by supposing that each user wishes to consume X at a standard quality and must pay to have any quality damage offset. So the cost of quality impairment to any individual is the amount he has to pay to undo the adverse effects of congestion. The greater the damage the greater must be the cost of offsetting it.

B. *Cost Functions.* Commodity Z is produced under competitive conditions of constant costs. Since we shall speak about the cost to each L^i of consuming a standard quality of X , we speak also of the cost to each L^i of consuming Z .

$$(1) \quad C_z^i = aZ_i \quad \text{so} \quad \frac{C_z^i}{Z_i} = a = P^z$$

where

C_z^i is the total cost to L^i of consuming Z_i units of Z

a is a constant

P^z is the price of Z

We treat Z as the numeraire good. Therefore, let $P^z = 1$.

Commodity X is also produced under conditions of constant cost, but congestion interferences add equally to the consumption costs for all L^i :

$$(2) \quad C_x^i = C_x(X_i, Q)$$

where C_x^i is the total cost to L^i of consuming X_i units of X

Q is the degree of generic congestion

$$(3) \quad Q = \frac{1}{K} \sum_{j=1}^3 \sum_{i \in L_j} W_j X_{ji} = \frac{T}{K}$$

where

K is the quantity of the assimilating medium—the capacity of the public good associated with consumption of X

X_{ji} is the amount of X consumed by $L^i \in L_j$

W_j is the interactive disturbance created by each unit of consumption of X by any member of L_j

T is the total interactive disturbance by the whole population

(3) shows that the size of congestion externalities depends on (a) total X , (b) the distribution of X among L_1, L_2, L_3 , (c) the capacity of the assimilative medium. Notice that each L^i adversely affects himself and others, the relative amounts depending on his membership in one of the polluter groups.

Some properties of C^i and Q (we henceforth omit subscript X in C_x^i): $\partial C^i / \partial Q$ is nonlinear:

$$(a) \quad \frac{\partial C^i}{\partial Q} = 0 \quad \text{for } Q \leq Q_0, \quad (b) \quad \frac{\partial^2 C^i}{\partial Q^2} > 0 \quad \text{for } Q > Q_0 \\ > 0 \quad \text{for } Q > Q_0.$$

where Q_0 is the threshold value of Q beyond which congestion effects occur.

K can be influenced by explicit public investment: social assimilation investment: I (for example, highway width or general sewerage treatment plants). Where $I=0$, $K=K_N$, the medium capacity given by nature (e.g., air and water as original "free goods" or unimproved earth surface for transportation). I is shown as annual costs—as though they represent either a nondurable resource use requiring annual replacement or the annual carrying charge of a durable investment. By this treatment we can compare annual cost with the annual value of the services rendered by the investment.

$$(4) \quad K = K_N + K(I)$$

W_j is the interactive disturbance rate specific to each L_j . Its value can be affected, however, by specific investments undertaken by the particular pollution groups: group treatment investment: I_j . These differ from social assimilative investments in being focused on the particular externalities generated by each member of the group—e.g., like smoke or fluid effluent recycling by each plant—rather than being facilities that deal with an aggregate of interactive effects from various sources.

$$(5) \quad W_j = W(j, I_j)$$

$$(a) \quad W_1(I_1 = 0) < W_2(I_2 = 0) < W_3(I_3 = 0)$$

$$(b) \quad \frac{\partial W_j}{\partial I_j} < 0 \quad (c) \quad \left| \frac{\partial^2 W_j}{\partial I_j^2} \right| < 0$$

Group investment decreases its own rate of pollution, but this is a decreasing effect. No I_j totally wipes out W_j .

C. *Demand Functions.* In order to isolate congestion issues, we assume that all L^i have the same income $\bar{Y}(t)$ and the same tastes for X and Z . Then, since every $L^i \in L_j$ has the same income, tastes and, by (3), faces the same cost of consuming X , $X_{ij} = \bar{X}_j$ (where \bar{X}_j is per capita consumption of X in group L_j). The demand for X and Z in L_j are:

$$(6) \quad \bar{X}_j = X \left(\frac{\bar{Y}}{P^*}, \frac{P_j}{P^*} \right) \quad \text{or, since} \quad P^* = 1, \quad \bar{X}_j = X(\bar{Y}, P_j)$$

$$(7) \quad \bar{Z}_j = Z \left(\frac{\bar{Y}}{P^*}, \frac{P_j}{P^*} \right) \quad \text{or} \quad \bar{Z}_j = Z(\bar{Y}, P_j)$$

where

P^* is the market price of Z to everyone

P_j is the market price of X to members of L_j

$$\frac{\partial \bar{X}_j}{\partial P_j} < 0 \quad \frac{\partial \bar{Z}_j}{\partial P_j} > 0 \quad \frac{\partial \bar{X}_j}{\partial \bar{Y}} > 0 < \frac{\partial \bar{Z}_j}{\partial \bar{Y}}$$

In the present model we shall subsequently examine the consequences of having government impose a congestion charge on each group per unit of X consumed by it. This congestion charge, Ψ_j , supplements the group's own cost with those overall system costs resulting from its contribution to congestion by an amount necessary to make the price facing each group equal the marginal resource cost to the whole population resulting from the group's incremental consumption of X , plus whatever group investments have been undertaken to decrease its congestion impact.

$$P_j = \frac{C_j}{\bar{X}_j} + \Psi_j + I_j \quad \text{where} \quad C_j = C^i \in L_j$$

$$(8) \quad \Psi_j = \delta C_j - \frac{C_j}{\bar{X}_j} \left(\delta C_j = \frac{\partial \left(\sum_k C^k \right)}{\partial \bar{X}_j} \right)$$

D. *Total Demand: The Effect of Population Growth.*

$$(9) \quad L(t) = L_0 e^{rt}$$

where

$L(t)$ is the total population at time t

L_0 is the initial population at time t

r is the rate of growth.

Since we assume the relative distribution among L_1, L_2, L_3 constant and in fractions n_1, n_2, n_3 :

$$(10) \quad \begin{aligned} N_j(t) &= n_j L_0 e^{rt} \quad j = 1, 2, 3 \\ \sum_{j=1}^3 n_j &= 1 \quad \left| N_j(t) \text{ is the number of members of } L_j \text{ at time } t \right. \end{aligned}$$

Then the total demand for X by L_j is:

$$(11) \quad \sum_{i \in L_j} X_{jit}(t) = \bar{X}_j(t) N_j(t)$$

And total X demanded at t :

$$(12) \quad X(t) = \sum_j \bar{X}_j(t) N_j(t)$$

Consequently $T(t)$ is given by:

$$(13) \quad T(t) = \sum_j W_j \bar{X}_j(t) n_j L_0 e^{rt} = L_0 e^{rt} \sum_j W_j \bar{X}_j(t) n_j$$

Since $\bar{X}_j(t) < \bar{X}_j(t_0)$ only if $T(t) > T(t_0)$, then, as population increases over time so does total interactive disturbance and congestion: $T'(t) > 0 < \dot{Q}(t)$. (Of course, since as a result prices rise, X rises slower than population growth alone would warrant.)

E. *Total Demand: The Effect of Income Growth.* As with population, let us examine the consequence of an upward drift of per capita income.

$$(14) \quad \bar{Y}(t) = \bar{Y}_0 e^{rt} \quad (\text{similar definitions as for (9)})$$

So \bar{X}_j is a function of time because of both population and income growth. Assuming $W_j = 0$,

$$(15) \quad \dot{t}(t) = \sum_j \frac{\partial T}{\partial \bar{X}_j} \frac{\partial \bar{X}_j}{\partial t} + \sum_j \frac{\partial T}{\partial N_j} N_j$$

$$(16) \quad \dot{\bar{X}}_j = \frac{\partial \bar{X}_j}{\partial \bar{Y}} \dot{\bar{Y}} + \frac{\partial \bar{X}_j}{\partial (P_j)} \dot{P}_j$$

\bar{X}_j tends to increase over time as per capita income rises, but tends to decrease only as P_j increases due to growing congestion. Thus, a for-

tiori, $T'(t) > 0$. Thus, $\dot{Q}(t) > 0$ also. As a result, $\bar{X}_j(t)$ will rise also, but due to the growing congestion, will rise slower than warranted by the pure income effect, especially as congestion becomes more and more serious.

This indicates that a system which begins with no congestion problem—with air and water and beaches uncongested, unpolluted free goods—will gradually move toward greater and greater congestion of all such natural media by the sheer growth and affluence of the society, especially when these occur in a context of greater spatial concentrations (urbanization) (since such growing concentration increases each W_j as it increases interdependence). The problem creeps up on the society that is doing nothing differently, only more and better—creeps, and then gallops.

F. *Optimal Resource Allocation between X and Z.*

1. *Optimal Conditions in Exchange and Production.* For each $L^i \in L_j$, an optimal budgetary allocation between X and Z requires:

$$(17) \quad \frac{P_j}{P^*} = MRS^i, \text{ or } P_j = MRS^i$$

where MRS^i is L^i 's marginal rate of substitution between X and Z . For all i, k , optimal production-consumption allocation requires:

$$(18) \quad MRS^i = MRS^k = \frac{MC_z}{MC_x} = \frac{MC_x}{a}$$

(since average cost of Z is an invariant a , so, too is marginal cost) MC_x is the marginal cost of consuming X (at standard quality). In our model this includes the various disposal, queuing and other interactive interference procedures that characterize congestion. The extent of these procedures determines the terms on which everybody can enjoy his own X , since it helps determine C^i . But these terms depend on who gets the additional (marginal) X . The three pollution groups differ in their efficiency in handling X . Group L_1 is the most efficient, L_2 next and L_3 least: $\delta C_1 < \delta C_2 < \delta C_3$. If all groups face the same price for X , intra-industry optimality requires that all of it go to L_1 , and none to L_2 or L_3 .

2. *Efficiency under Zero Congestion Charges.* Assume that $\Psi_j = 0$, for all j . Then there is a high probability that no group will find it worthwhile to make a group treatment investment. Since most of the impact of each group's interactive disturbance is on others rather than on itself, any such investment largely benefits those others. Profitability for self-interest would require that a very small portion of the project's overall benefits be large enough to exceed the whole cost of the

investment. Benefit to the investor from such a project is given as:

$$(19) \quad \Pi_j^i = -\frac{\partial C^i}{\partial I_j} = -\left[\frac{\frac{\partial C^i}{\partial Q} X_j}{K} \frac{\partial W_j}{\partial I_j} \right] + \left[1 - \frac{\frac{\partial C^i}{\partial Q} W_j}{K X_j} \frac{\partial X_j}{\partial \left(\frac{C^i}{X_j} \right)} \right]$$

Benefits to the population as a whole is given as:

$$(20) \quad \Pi_j = L_j \pi^i \Pi_j^i$$

With this disparity it is most doubtful that there exists an I_j^i such that $\Pi_j^i \geq I_j^i$. Thus, with $\Psi_j = 0$ we can expect also that $I_j = 0$ for all j .

If this is so, then all individuals face the same $P_j = C^i/X_j$. Consequently, all have the same incentive to consume any marginal resource transfer from Z ; i.e., all will bid equally to consume an extra unit of X . As a result, any marginal resources flowing to X are likely to be divided among the three groups in the same proportions as their percentage of the population: n_1, n_2, n_3 . So the average of the marginal costs of consuming X that follows a marginal resource transfer from Z is $\sum_j n_j \delta C_j$, which exceeds $\min(\delta C_1, \delta C_2, \delta C_3)$ (where $\delta C_j = \sum_k \partial C^k / \partial X_j$). This has two implications.

First, it means that whatever resources flow from Z to X are inefficiently distributed within X (as noted in the last section). Members of L_2 and L_3 consume too much relative to L_1 . Second, when L^i experiences incremental consumption he pays $\partial C^i / \partial X_i$ (marginal private cost). But his extra consumption adds to everyone else's cost as well: $\partial \sum_{k \neq i} C^k / \partial X_i$ (i.e., marginal social cost less $\partial C^i / \partial X_i$). Thus, the marginal social cost substantially exceeds the marginal private cost: $\delta C_j > \partial C^i / \partial X_j$. Consequently, resource flow between Z and X is determined by an aggregate demand for X which is systematically biased upward above the socially optimal amount, because each buyer faces a price lower than the marginal social cost of supplying it. Too much X is consumed, too little Z .

Optimal allocation would be arrived at where each $L^i \in L_j$ were faced with the true δC_j as his cost of consuming extra X . Then members of L_1, L_2 and L_3 , faced with differentially higher prices of X in the same order, would cut their consumption differentially in that order. The lesser per capita consumption in L_2 and L_3 relative to L_1 would be

determined at those points where the resulting higher individual marginal valuation of X equalled the differentially higher prices.

By comparison, the suboptimal situation shows that there is both a distributional and an allocational distortion. The distribution issue underscores what might be called the customary pollution aspect: heavy polluters making too much use of a scarce medium relative to light polluters (or victims). The allocational issue underscores the customary congestion aspect: everyone's activity damaging the quality of others' and his own prospects.

3. *Congestion Charges.* Let us now examine the consequences of having the government impose congestion charges on everyone so that each L^i pay a price for consumption of X which reflects the true marginal social cost of that consumption. The total revenues collected are subsequently redistributed equally among the same L^i . Since each L^i is homogeneous, charges will differ only with respect to group. The charge Ψ_j is set so that each L^i pays $P_j = C^i/X_j + \Psi_j + I_j$ and

$$(21) \quad \Psi_j = \delta C_j - \frac{C^i}{X_j} = \sum_k \frac{W_j}{K} \frac{\partial C^k}{\partial Q} - \frac{C^i}{X_j}$$

(from (2) and (3))

The result of this has been discussed in the last section. Every L^i faces the true marginal social cost of his consumption of X . His relative demand between X and Z is undistorted. So aggregate demand for X is undistorted. Less X will be consumed, more Z , and the distribution of X will show $\bar{X}_1 > \bar{X}_2 > \bar{X}_3$, because, by (21), $P_1 < P_2 < P_3$. Since generic congestion in our model is shared equally by all, the relative price changes among L_1 , L_2 and L_3 result in a real income redistribution. All are benefited by an increase in Z and a decrease in congestion cost associated with the total consumption of X . But L_3 pays most per capita for these benefits in its members' sacrifice of X consumption, L_2 pays next most, and L_1 least, for the same benefits. At the least there is a change in relative income distribution. But there is likely to be an absolute change too. Members of L_3 may actually be worse off after the resource shift than before, since they pay much more in charges than they receive in per capita refund. Members of L_2 may gain or lose on balance from the shift, but probably not by much in either direction (depending, of course, on the technological, taste and relative numbers characteristics of the actual situation). If per capita refund is nearly as great as their average charge, they will gain because of

the lower congestion cost in X . Members of L_1 probably gain substantially—both because of a refund which greatly exceeds the charge, and because of the lessened congestion in X : gains which, by the compensation principle, are more than sufficient to make it possible to pay off the losers to make them no worse off than before. Congestion charges damage the "polluters" and help the "victims."

G. *Group Treatment Investment.* When $\Psi_j = 0$ for all there is little incentive for individuals to undertake group treatment investment to lessen the externality effect of their consumption. All I_j is likely to be zero. But imposition of congestion charges changes this. Ψ_j is a positive function of W_j (by (21)), and the size of W_j can be decreased by such investments. Ψ_j and I_j are substitute outlays, then, and each individual will undertake such investment as decreases Ψ_j by more than the cost of the investment. In effect, he determines his optimal I_j by minimizing P_j with respect to I_j . I_j represents a "treatment cost" per unit of X consumed by each member of L_j . So it enters the unit price of X facing each such member:

$$(22) \quad P_j = \delta C_j + I_j$$

The condition for the optimal I_j is therefore given as:

$$(23) \quad \frac{1}{K} \sum_k \frac{\partial C^k}{\partial Q} \frac{dW_j}{dI_j} = -1 \quad \text{or} \quad \frac{dW_j}{dI_j} = - \frac{K}{\sum_k \frac{\partial C^k}{\partial Q}}$$

The left-hand term in the first version represents the impact on marginal social costs of a marginal investment. Investment is more favored the higher its impact on W_j , the more strongly increments in congestion increase the necessary offset costs to achieve standard consumption quality, the larger the overall population, or the lesser the assimilative capacity of the common medium. The first is a technological consideration. The others essentially denote the several elements comprising the seriousness of the congestion problem. Investment is favored the more serious is current congestion and the more effectively it can decrease the per unit contributions to that condition. It should be noted that each I_j is a substitute for the only other policy variable implied in (23), I (social assimilation investment), since the latter increases K and therefore decreases the payoff to each I_j .

The only term which can reveal which groups are likely to invest is dW_j/dI_j . This is subject to technological possibilities. We may surmise, however, some properties of the relationship. The

larger is W_j , the easier it probably is to effect a unit decrease in it, since critical mass phenomena may make several types of recycling economical or may make available a greater variety of ameliorative procedures. Thus, we assume

$$(24) \quad \frac{dW_j}{dI_j|_{I \rightarrow}} > \frac{dW_k}{dI_k|_{I \rightarrow}} \quad \text{if } W_j(I_j = 0) > W_k(I_k = 0)$$

Since $W_3 > W_2 > W_1$, this suggests that greater polluters are more likely to substitute treatment investment for congestion penalties than lesser polluters. It is even conceivable that the treatment will be carried far enough to wipe out the difference in observed W_j between two adjacent groups (the group will, however, remain distinct because the originally higher group will have achieved the lower W_j only by a greater I_j than that of the lower group). Whatever the actual pattern of I_j among groups, the presence of the congestion charges (in the context of equal incomes, tastes and congestion sharing) will guarantee that the total resources spent on treatment will be efficiently spent: it will buy the largest total decrease in congestion cost for the population; i.e.,

$$(25) \quad \sum_j \frac{\partial C_j}{\partial I_j} dI_j N_j - \sum_j dI_j N_j = \Pi(I_{Z_j}^*) \max \text{ for } \sum_j dI_j N_j = I_{Z_j}^*$$

where $\pi_j(I_{Z_j}^*)$ is the aggregate cost improvement resulting from a given total of $I_{Z_j}^*$ in resources being used for I_j by individual users as part of their utility maximizing decisions, less the amount $I_{Z_j}^*$; i.e., an aggregate private investment profit.

H. Social Assimilation Investment. Social assimilation investment (I), unlike the private projects of I_j , is carried out by the collectivity of L —the government. It decreases the overall system costs of congestion by increasing the assimilative capacity of the common medium which is shared by users of X . Multiple user treatment plants, road capacity enlargement, and low-flow augmentation are examples of this type of investment. We have already noted that I is a substitute for I_j . It is also a substitute for more production of X or Z . Its justification therefore depends upon traditional cost-benefit analysis. This consists in comparing the aggregate cost savings from decreased congestion less the loss of net cost saving from private treatment investment that is precluded by the public investment—the project's benefits—with the project's costs (i.e., the oppor-

tunity cost in terms of Z). The marginal profit from an increment of such investment is:

$$(26) \quad \frac{\partial \Pi}{\partial I} = \left[\frac{\partial \sum_k C^k}{\partial I} - \frac{\partial (\Pi_j)}{\partial I} \right] - 1 \\ = - \frac{T}{K^*} \frac{dK}{dI} \frac{\partial \sum_k C^k}{\partial Q} - \frac{\partial (\Pi_j)}{\partial I} - 1$$

The second term is complex, since it shows the effect of a changing K on the cost impact of each marginal I_j , and the effect of this on the optimal I_j . It is therefore difficult to characterize the properties of $\partial \Pi / \partial I$ in general. More precise specification of the components is necessary for deeper analysis. With a precise specification we can, by setting $\partial \Pi / \partial I = 0$, find the optimal $I - \hat{I}$ —and for each user thereby, the optimal set of private investments, $\hat{I}_1, \hat{I}_2, \hat{I}_3$. This in turn establishes the set (W_j) and so the set (Ψ_j). All of these together determine the marginal social costs for different users of X , and for X as a whole relative to Z . Thus, the overall allocation of resources and distribution of income are determined.

The most obvious qualitative result suggested by (23) and (26) is an asymmetry between I and the set (I_j). The former is, but the latter are not, a positive function of T (aggregate interactive disturbance). Since $\hat{T}(t) > 0$, then we can expect also: (27) $\hat{I}(t) > 0$. I will be increasingly justified by population and income growth. The profitability of each I_j will not be directly enhanced by a growing T , but to whatever extent this growth in T stems from population growth there will be more total users, each one of whom has an unchanged incentive for I_j . However, the growing I and thereby the ever increasing K will systematically decrease the attractiveness of private treatment expenditures. Growth may well tend to favor a gradual switch from private to public investment, assuming that decreasing returns to I (i.e., the size of dK/dI) do not set in substantially. What could mitigate this trend is for incremental public investment to have so disastrous an effect upon the profitability of private investment that benefit-cost considerations justify only small increases relative to population growth (which tends to increase $\sum_j I_j N_j$ —total private investment). But precise characteristics of (26) and the set (23) are necessary to determine this.

One subsidiary result is that public investment, unlike private, does not require the presence of congestion charges to justify its existence. As with the latter, the presence of charges affects the opportunity costs of undertaking investment. But

the direction of effect is reversed. Congestion charges, by making private investment attractive, make the opportunity cost of public investment higher, because public investment tends to displace private. In the absence of charges there would be no private investment and, therefore, none to be displaced by public. Public investment would be considerably larger. The resulting intensity of congestion is difficult to compare under the two situations. It will generally depend on the effectiveness of individuals versus multi-user methods of controlling congestion. A rough guess is that some forms of congestion are easier to control individually than collectively, so the situation that can find room for both—i.e., with charges—will use resources more efficiently in this regard.

I. *Conclusion.* Only the briefest conclusion can be given here. A system with congestion charges will differ from one without in a variety of dimen-

sions. Relative prices between X and Z , and among different users, will be affected. But so too will be the amount of private and public investment to control congestion. While no definitive propositions can be made at this level of generality once the complex interaction of all these dimensions is allowed, the no-charge system is likely to display the following characteristics relative to the more efficient charge system: (1) too much consumption of X relative to Z ; (2) too much consumption of X by L_3 , especially and L_2 relative to L_1 ; (3) not enough private investment and hence too high congestion generation rates per unit of private activity; (4) too much relative dependence on public investment to carry the burden of control against a growing congestion problem; and (5) in sum, more congestion at any time, and a more rapid worsening of the problem over time.

FOREIGN ENTERPRISES AND DEVELOPING NATIONS IN THE RAW MATERIALS INDUSTRIES*

By RAYMOND VERNON

Harvard University

The object of this paper is to review the basis for the struggle between foreign investors in raw materials and host governments in less developed countries and to consider what its next phase is likely to be. In these countries, most foreign investment is to be found in a relatively small number of materials: petroleum, copper, iron ore, and bauxite, among the minerals; rubber, bananas, tea and timber, in the field of agriculture.¹ Though our analysis is general in application, it concentrates on the cases of oil, copper, and bauxite.²

The Case of Oil

In the international oil industry, as in so many other industries in which multinational enterprise dominates, the story begins with the U.S. domestic market. By 1860, within a decade of the industry's creation, the leaders in the industry had extended their business from the U.S. domestic market to the export trade. By 1870, two-thirds of U.S. production was being exported. For years afterwards, the international oil trade was dominated by the exports of U.S. firms, challenged only by the Russians and the Dutch.

The interest of U.S. oil companies in foreign markets was followed eventually by a strong interest in foreign sources of oil as well. That interest arose in part from the characteristic oligopoly structure of the industry, generating a typical sequence of "imperfect" competitive responses. Those responses have changed very little over the decades.

As the oil companies see it, there are important scale economies available in crude oil exploitation and distribution, some a consequence of the insurance principle, some a function of high fixed costs. With declining costs, each of the suppliers of oil is hesitant to use price reductions as a means of stabilizing its production and sales, fearing off-

setting reductions from the other participants. To secure stability of markets, therefore, each producer tries to develop its own downstream refining and marketing facilities. At the same time, however, those refiners and marketers that lack their own sources of material supply are fearful that the integrated oligopoly sellers are in a position to wipe out the profit margin by raising the prices of the raw materials. The market-oriented operators, therefore, try to capture sources of crude oil in order to reduce their vulnerability.

The U.S. oil exporters made their early investments in foreign oil production to establish a source for their export markets in Asia and the Mediterranean basin.³ The Dutch, because of their access to Indonesian oil, seemed capable of invading the Asian markets; the Russians posed another threat by reason of their overland access to European markets; so did the British, who by 1900 had begun to develop sources in the Middle East. In general, the major oil companies had to take an interest in any potential source of oil, wherever there was a risk that the source, when developed, might undersell existing supplies.

It would be oversimplifying the situation to think of such investment merely as an effort to maximize profit yields at the margin. In economic terms, the cost of development could better be attributed to the hedging of risk—the risk of losing control of the price structure in established markets. The expenditure also could be thought of, in part, as an investment—an investment in the acquisition of knowledge about local production costs.⁴

* For brief accounts of the early discoveries in oil, see E. T. Penrose, *The Large International Firm in Developing Countries* (London: George Allen & Unwin, Ltd.), pp. 53–62; also Staff Report to the Federal Trade Commission, *The International Petroleum Cartel*, U.S. Senate, 82nd Cong., 2nd sess., Select Committee on Small Business, Subcommittee on Monopoly, Committee Print No. 6 (Washington: G.P.O., 1952), p. 37 *passim*.

⁴ An even more complex motivation for investment, akin to the idea of the effective deterrent in game theory, is found in the official Royal Dutch history: "Standard Oil had obtained a firm footing in the Netherlands Indies. . . . Being hard pressed by this keen competition, the Royal Dutch/Shell group made a successful attempt to extend its business to the American continent." *The Royal Dutch Petroleum Company 1890–1950* (The Hague: 1950), p. 18.

* The preparation of this article was financed by grants from the Ford Foundation to the Harvard Business School and to the Harvard Center for International Affairs.

¹ United Nations Conference on Trade and Development, *UNCTAD Commodity Survey 1966* (New York: United Nations, 1966), pp. 66, 67, 85.

² Perhaps the best account of U.S. foreign direct investment during these early years is that of Mira Wilkins, *The Emergence of Multinational Enterprise*, to be published by the Harvard Univ. Press in 1970.

In addition, the oil companies soon discovered that a well-diversified supply of resources was especially useful in dealing with blockages of supply, whether threatened or actual. This lesson was borne home repeatedly over the decades: in the bitter negotiations with Mexico in the 1930's; in the battles with Mossadegh over Iran's oil in the early 1950's; in the various closings of the Suez Canal in the 1950's and 1960's.

As a result, the number of foreign installations of U.S. parent firms grew persistently until World War II and explosively thereafter. Counting only the nine largest U.S.-controlled enterprises, one finds the number of overseas subsidiaries rising from only 102 in 1920 to over 1,400 in 1967; and of these 1,400, about 30 percent were located in the less developed areas. With the spread of the big international companies, the old geographical spheres of dominance associated with each company were badly eroded. Although there had always been a certain amount of market interpenetration among the big companies, the degree of such interpenetration increased. While the large U.S. firms managed to secure refining and distributing positions in European markets, the British and Dutch companies succeeded in expanding their foothold in North American markets. Today, although each major company is still associated with an explicit national identity, the spheres of interest of these companies no longer can be defined on simple geographical or political lines.

Although the major firms have been extending their reach into one another's territory, however, the industry has not grown more concentrated in structure. Quite the contrary. The expansion of world markets since World War II has brought a new crop of actors into the international oil business. This development has occurred because the barriers to entry into the petroleum business seem to have declined. There has been a gradual dissemination of the necessary exploration and refining technology and a gradual increase in the number of firms capable of assembling the money and organization appropriate to the activity. Moreover, the characteristic oligopoly responses to threat have crossed industry lines. As the oil companies have moved "downstream" into chemical manufacturing, some of the chemical companies have moved "upstream" to secure their own sources of crude oil. As a result, the tight and tidy industry structure of the 1928-48 period has grown more diffuse and more complex.

At the same time, the role of state-owned enterprises as sellers of oil has grown. A number of state-owned oil companies existed prior to World

War II, to be sure. But those companies had generally been set up by countries that were concerned about oil imports, not exports. The state-owned oil companies created since World War II, however, have been set up, not only by importing countries such as India and Thailand, but even more importantly by exporting countries such as Venezuela and Iran. This new crop of companies, like their predecessors, is eager to control the refineries, transport facilities, and markets on which their security depends. And some are managing to take a few first steps toward their objective. But as far as the oil exporters are concerned, their aspirations will exceed their reach for some time; it is no easy task for outsiders to acquire the necessary distribution facilities in other countries. How to accommodate these aspirations to world oil markets is a major problem confronting the industry.

Copper and Aluminum

Many of the major tendencies noted in connection with the international oil industry are to be found in the other major raw material industries as well.

U.S. enterprise has been prominent in practically all major nonferrous metal industries. As with oil, that prominence was characteristically built up in the first instance on the basis of a strong domestic market. This has generally been followed by the emergence of a vigorous export trade, by a persistent drive toward vertical integration, and by a steady geographical spread.

A global count of subsidiaries records the swift spread in the interests of U.S. firms. Between 1919 and 1967, the number of foreign subsidiaries of nine major U.S.-controlled raw material enterprises increased from 17 to 417; and of these 417, as much as 189 were settled in the less developed areas. Moreover, extraction activities spread much slower than activities in manufacturing and sales. Indeed, for our group of nine U.S. enterprises, there is only slow growth in the number of extractive operations between 1957 and 1967, reflecting an absolute decline in a few countries such as Mexico and slight rises elsewhere. Thereby hangs a tale, to which we shall shortly return.

In all these developments, there were differences worth noting among the industries. One such difference lay in their seeming bargaining strength. The aluminum industry, for instance, gave off an aura of invulnerability in its relations with the less developed countries such as was associated with the oil industry thirty or forty years ago. The copper industry, on the other

hand, seemed much more vulnerable in these relations.

Perhaps the apparent difference in the vulnerability of foreigners in the two industries arose in part from differences in their starting condition. From the first, the main barrier to entry in the copper industry was at the mining stage, whereas the main barrier to entry for aluminum was at the refining stage; aluminum producers therefore bargained with foreign governments from a position of greater strength.⁶ Besides, although both industries are concentrated at the primary and crude fabricating level, the aluminum industry has been the more highly concentrated; it has demanded larger installations and much more capital intensive facilities.⁶

Moreover, there is the question of the role of technology. In technological terms, aluminum still has the earmarks of a young industry, with a rapid rate of technological change at every level, from reducing and refining to processing and fabricating. The copper industry seems to present a rather different picture. In that industry, the technology of refining and fabricating appears to be changing less rapidly and to be more easily available to outsiders.⁷

On all counts, the copper industry would be adjudged more vulnerable to pressure, as seen from the viewpoint of the nations that control the raw materials. We explore that viewpoint next.

The Urge for Control

The record of recent decades has been an unceasing effort on the part of less developed countries to extend their control over the exploitation of the raw materials that originate in their borders.

In the earliest stages of raw material investments, of course, the issue of control generally was quite unimportant. Countries that have little prior experience with raw material investment by foreigners generally look upon the first discoveries as a windfall—a windfall whose exploitation is totally beyond the capability of the host country

itself. An attractive short-term strategy for the less developed country in these circumstances is to invite as many foreign companies as possible to bid for the putative resources and to turn over the job of exploration and exploitation to the highest bidder. As a general rule, less developed countries have pursued that strategy in the early stages.

The first bone of contention between the foreign investor and the host government has been the question of division of the risks and rewards rather than that of control. Over time, the skill of host countries in securing greater rewards and assuming lesser risks has steadily increased. For instance, Libya's handling of her oil negotiations in the 1950's, aided by the advice and counsel of British industry experts, was a classic in the protection of Libyan interests.

The increasing capacity of host governments to reduce their risks and increase their rewards has been abetted by another factor, one of even greater importance. Once a raw material project is successfully launched—once the material has been located, assayed, lifted out of the ground, transported, and shipped—the perceived level of risk associated with the enterprise declines precipitately. The returns to the foreign company no longer seem appropriate to the risk, and the government feels justified in demanding downward adjustments in the investor's share of profits.⁸ Meanwhile, the disposition of the investor to accede to marginal pressures, if this will protect the investment, tends to grow.

The success of governments in elevating their share of the profits is illustrated by the data in Table 1. Though the data are subject to various statistical weaknesses, as any "share of profit" figure is bound to be, they serve well enough to indicate the trend.

Governmental efforts constantly to raise their share of the profits from raw material exploitation can be cited as a near-immutable law of governmental behavior. The problem for governments, however, is not only one of share but also one of absolute level. A supply of government funds creates its own long-term demand. Accordingly, once governments have experienced a given level of revenue, the risk of interruption becomes intolerable to bear. A few figures are helpful in suggesting the magnitude of the problem from the

⁶ For insights into these differences, see M. S. Brown and John Butler, *The Production, Marketing, and Consumption of Copper and Aluminum* (Praeger, 1968).

⁶ The generalization can be tested by value added per establishment and horsepower per establishment from the U.S. Census of Manufactures and by data from the financial statements of major firms.

⁷ OECD, *Gaps in Technology between Member Countries: Non-Ferrous Metals* (Paris, 1968), mimeo.; also M. J. Peck, "Inventions in the Postwar American Aluminum Industry," in R. R. Nelson, ed., *The Rate and Direction of Inventive Activity* (Princeton Univ. Press, 1962).

⁸ This point is carefully analyzed in R. F. Mikešell, "Conflict in Foreign Investor-Host Country Relations: A Preliminary Analysis," in R. F. Mikešell, ed., *Foreign Investment in the Petroleum and Mineral Industries: Case Studies on Investor-Host Relations*, to be published in 1970.

viewpoint of host governments. In 1964, for instance, the Venezuelan government seems to have derived something like 60 or 70 percent of its gross revenue from taxes paid by foreign petroleum and iron ore operations.⁹ In Iran in the fiscal year ending in 1966 the figure for oil operations was about 50 percent.¹⁰

There are, of course, other measures that mirror the sense of heavy dependence of the less developed countries on foreign raw material operations—measures, for instance, of their contribution to the gross product of the country and to the foreign exchange revenues of the country. All these measures point in the same general direction: an increase in the seeming vulnerability of the economy to changes in these foreign-controlled operations. The very success of the less developed countries in drawing the foreigners into their economy, therefore, has created new perils, both for the foreigner and for the country.

The Past and the Future

On the record, the foreigner who invests in raw materials exploitation in the less developed countries has no reason to hope for long-run tranquility and stability. The investments of foreigners in tropical plantations, popular during the late nineteenth and early twentieth century, have almost passed into history. The investments in minerals and petroleum are under constant pressure. But there appears to be system and order in the timing and effectiveness of the pressure. Those undertakings that have least to offer in the way of capital, technology, and markets appear to be most vulnerable.

As the governments of less developed countries sense a weakening in their need for the foreign investor, they can be counted on to press foreign investors for an increased share of the profits and an increased measure of control over the exploitation of raw materials. As success is achieved by the exporting countries, however, there is a risk that the victories may prove Pyrrhic in character—that the aggregate profits will shrink and that the “control” acquired by the countries will be illusory. Why?

The simple arithmetic of profit sharing sug-

⁹ Sources differ somewhat on the exact figure, presumably because of definitional differences. See W. G. Harris, “The Impact of the Petroleum Export Industry on the Pattern of Venezuela Economic Development,” and Henry Gomez, “Venezuela’s Iron Ore Industry,” both in R. F. Mikesell, ed., *op. cit.*

¹⁰ In addition to the chapters already cited in the R. F. Mikesell work, see particularly W. S. Bortsch, “The Impact of the Oil Industry on the Economy of Iran.”

TABLE 1
HOST COUNTRY SHARE OF PRE-TAX PROFITS OF
FOREIGN INVESTORS IN RAW
MATERIAL ENTERPRISES

	Venezuela (Oil)	Chile (Copper)
1930	*	16%
1940	58%†	28
1950	51	58
1955	52	69
1960	68	65
1965	66	69‡

* Not available. Probably less than 30 percent.

† 1943.

‡ 1964.

SOURCE: R. F. Mikesell, ed., *Foreign Investment in the Petroleum and Mineral Industries: Case Studies on Investor-Host Relations*, to be published in 1970. The Venezuela oil data are from a chapter by G. G. Edwards, “Foreign Petroleum Companies and the State in Venezuela,” while the Chile copper data are from Markos Mamalakis, “The American Copper Companies and the Chilean Government, 1920–1967.”

gests that the international firms will not accept much further reduction in their share of profits, certainly not at the rate at which those shares have been reduced in recent years. Until now, the international oil and copper firms have generally managed to make up through increased turnover or higher prices what they were losing through declining shares. Now that the producing countries have raised their share of the profit close to 70 percent, however, the arithmetic of the situation grows more difficult. If the companies’ share were to be reduced from 30 to 20 percent, for instance, sales would have to increase by as much as one-half in order to offset the effects of the declining share. Nevertheless, governments have a compelling political need to demonstrate to a domestic constituency their resoluteness, and their independence in negotiations with foreign investors. For this reason, governments will find it necessary to push their demands from time to time, even if the consequences seem threatening to their national well-being.

International raw material producing companies, therefore, may be pushed to the point at which their corporate strategy no longer assumes any substantial measure of control over the supply and price of raw materials that originate in the less developed world. That shift would end the community of interest which now generally exists between the raw material exporting countries

and the large international companies under which both benefit from high prices in raw material exports.¹¹ Once the companies had relinquished their hold, whether under duress or by choice, the next question for them to consider is whether their interests would be served by more open and competitive international markets for raw materials.

The possibility of such a shift in strategy, though plausible, would still be regarded as revolutionary by most large enterprises, long conditioned to a situation of control over their raw material sources. Moreover, they may see some risk that the producing countries will be able to find a basis for restricting supplies and raising prices.

Instead of withdrawing from the raw materials business, therefore, the international companies may choose the strategy of developing a greater degree of self-sufficiency based on the resources of the North American continent and Western Europe. In the case of oil, the vast new discoveries in Alaska and Canada could bring about that re-

sult willy-nilly. But the same trend could occur in some of the metals as well, especially iron ore and the copper-lead-zinc group; large integrated companies could turn to North America and Western Europe for their raw materials. How far the companies can go in this direction will depend partly on the degree of protection that the United States, Canada, and perhaps Western Europe were prepared to extend. The possibility of Western Europe's being included in any such protective arrangement is somewhat enhanced by the extensive mutuality of interest that has been developing among the major raw material producing companies of the advanced world, irrespective of nationality. The strategy that could well emerge, therefore, is one in which the advanced countries maintained a sheltered market for a portion of their domestic needs, leaving it to the less developed countries to supply the rest of those domestic needs as best they could.

The less developed countries that export raw materials, therefore, could end up swapping one form of dependence for another. As they terminate their organic ties to the large international companies, they may reduce their vulnerability to the internal logistical decisions of the large companies; but the effect of that achievement is to expose them much more acutely to the risks and vagaries of an international market in which arm's-length sellers compete. Which of these types of risk will prove the more tolerable for the less developed countries remains to be seen.

¹¹ Needless to say, the community of interest with the large international oil companies in maintaining high prices has been less evident for countries that are on an import basis for oil. See Michael Tanzer, *The Political Economy of International Oil and the Under-developed Countries* (Beacon Press, 1969), especially Chaps. 18 and 22. The argument there is that Indian and Mexican public investments in oil have been highly rewarding to those countries.

DISCUSSION

RICHARD R. NELSON: Rothenberg has developed an interesting and potentially fruitful conception of "generic congestion." He has proposed that pollution and congestion are similar in that they both are external diseconomies (given existing market arrangements) resulting from exploitation of a "public good" by certain activities. He has suggested that both pollution and congestion tend to increase more than proportionately to the increase in the activities that generate them with possibly a threshold level before there is any significant effect at all. Pollution and congestion are similar also in that each can be coped with by expanding the capacity of the public good in question, or by measures to reduce the external diseconomies generated by activities. Rothenberg then suggests that the key difference between congestion and pollution is that in congestion all of us contribute to the external diseconomy and all are hurt by it and in roughly the same magnitude and way. In pollution, on the other hand, one can distinguish between those that pollute and those that suffer from pollution, or at least one can identify significant differences among people in terms of cost generation, cost bearing, or both. My remarks will be concerned with exploring this last suggestion of Rothenberg.

But before proceeding let me remark that while I found the verbal theorizing clear and provocative I did not find that the mathematical treatment added anything at all, either in terms of sharpening and clarifying concepts or in terms of permitting one to see interesting implications that were not apparent from the verbal discussion. I wonder, therefore, what Rothenberg's purpose was in presenting the mathematics.

I found Rothenberg's distinction between congestion and pollution interesting because it seemed to provide a handle for the start of positive analysis of the ability of society to cope with these problems. I take it as evident that, initially at least and undoubtedly at the present time, the growth of congestion and pollution is associated with lack of Pareto optimality. There exists allocations of resources and side payments that can make everyone better off. I take it that the new positive theory of public goods (and bads) evolving from Coase, Demsetz, and Buchanan rests heavily on the argument that if mutually beneficial deals can be made, there are forces at work trying to make them. But there will be lags and difficulties because the institutional structure of the "trading" will not have the nice properties of a competitive market. There will be problems of bargaining and enforcing. If there are a large number of people involved, there will be the "free rider" problem. The group "contract" may be difficult to achieve. What intrigues me about the Rothenberg distinction between congestion and pollution is that, if he is right, it seemed to me at first that it would be easier to get agreed political action on the congestion problem than the pollution problem. For congestion Pareto optimality requires roughly the same changes in everyone's behavior (under Rothenberg's definition). Everyone is a net beneficiary from a change that affects everybody in roughly the same way. No side payments of complicated distinctions

are required. For pollution the required changes affect different people and groups differently. There are conflicting interests. To resolve these means that some interests have to be overridden or a complicated set of side payments or logrolls developed.

Of course, this argument rests, not only on some untested and vague hypotheses about the nature of political processes, but also on the presumption that Rothenberg's distinction between congestion and pollution is roughly consonant with what we normally mean by these terms. Unfortunately, in many important cases Rothenberg's distinctions do not appear to correspond to general language usage. For example, crowding of highways by automobiles does fit (at least roughly) the Rothenberg characterization of congestion. But then so does the contamination of the atmosphere by automobile exhaust. Yet the latter we tend to call pollution, not congestion. I would suggest that, at least to members of the Sierra Club (who were there first), the crowding up of the park areas by weekend picknickers from the city fits the Rothenberg characterization of pollution more than congestion.

Let us consider political action on a couple of important kinds of congestion or pollution. We certainly have had a lot of action (investment) to deal with highway crowding. Here the hypothesis that if "we all can benefit if we all pay a bit into the kitty," then "we will all pay into the kitty," together with the supplementary assertion that this kind of congestion is like Rothenberg's definition of congestion, comes out pretty well. That we have had much more political hassling about public investment and conservation rules regarding outdoor recreation areas I think testifies that, as hinted above, this case does not fit Rothenberg's definition of a congestion case. People differ greatly in the extent to which they feel they bear the costs of this kind of congestion. Some people even like crowds and feel a little lost and lonely with no one about. One would expect this kind of congestion or pollution problem (which is it?) would be a harder nut to crack politically.

One interesting question is, if the arguments above are correct, why have we made so little headway towards dealing with the automobile smog problem which seems to have roughly the same structure as the automobile highway crowding problem. Of course, the political process hypothesis is very crude. One might pose another that if there is a small minority that "causes the trouble" and everyone else bears the costs, the political process can move relatively rapidly to deal with the problem. Thus we seem to do better against industrial pollution than automobile pollution. But why are we having so much trouble with the latter? People have proven willing to pay highway taxes, why not smog taxes?

I think Rothenberg's distinction between congestion and pollution raises as many questions as it answers, and my suggestions about the nature of "hard" and "easy" public policy problems are of the same breed. But I think the most important questions in the study of problems like congestion and pollution (and poverty

and central city problems) involve the nature of the public policy processes involved. In an economy where there is so much concern with public goods and bads one cannot rest content with a theoretical structure that makes positive assertions only about the determinants of private goods and bads. Nor can we feel our job as economists is done when we characterize Pareto optimal or at least "better" configurations, as does Rothenberg. We have to develop a workable theory of nonmarket decision making.

CHANDLER MORSE: The papers by Vernon and Christy make an interesting and somewhat complementary pair. Vernon notes the increasingly effective efforts of the raw material producing countries to broaden their shares of the economic rents hitherto garnered by foreign corporate enterprises, and to share also in the control of local operations. According to Vernon, corporate managements must either go along with these developments or turn increasingly to reliance on the high-cost natural resources of the Western World or possibly on not-so-high-cost synthetics. Christy, on the other hand, points out that a third option is rapidly emerging; namely, to exploit the organic and mineral resources of the open seas. Vernon suggests that if the multinational companies choose the second option, their advanced technology and the ability of their home governments to protect the principal markets for raw materials means that the raw material producers may be left holding a bag full of decreasingly valuable natural assets for which they can find neither uses nor buyers. Christy, in turn, implies that in the coming struggle for control over whatever it turns out the oceans can provide, victory also seems likely to go to the economically and technically advanced countries. Since he apparently dislikes both the distributional implications of this possibility and the probabilities of conflict arising from efforts of the many and weak to prevent monopolization by the few and strong, he makes a plea for mobilization of the social-scientific arts in an effort to develop a more satisfactory alternative. In effect, he is calling for rationalizing and humanizing the third option.

The necessity for action arises because technological changes and other factors are making the tradition of "freedom of the seas"—of communal rights in the ocean's wealth—increasingly unsatisfactory. The situation is remarkably close to that hypothesized by Harold Demsetz in a paper, "Toward a Theory of Property Rights," presented at these meetings three years ago. (*A.E.R.*, May, 1967.) In an effort to explain the emergence of private, or exclusive, property rights, Demsetz presented the hypothesis that "property rights develop to internalize externalities when the gains of internationalization become larger than the cost of internalization." This, it seems clear, is exactly what Christy says is happening, so far as circumstances permit, with respect to the tradition of open access on the high seas. However, it is happening in a unilateral and arbitrary manner that is likely to have heavily one-sided distributive effects unacceptable to those who lose out. To understand why this is so and some of the problems of

developing an alternative, we may again turn to Demsetz.

The essence of the problem, according to his analysis, is the relation of the gains to be derived from internalization of benefits and costs to the costs of achieving such internalization. In the present case, we shall take the gains for granted, for unless they existed there would be no problem. Their identification, classification, and quantification are important but beyond the scope of this comment.

What about the costs? Demsetz divides the costs of internalization into two categories: (1) costs of negotiation; i.e., of reaching agreement concerning who shall enjoy what rights to exclude whom; (2) costs of policing; i.e., of effectively maintaining whatever exclusivity may be established, whether by agreement or otherwise.

For present purposes it will be useful to broaden the first category into something we shall call, "costs of establishing exclusive title," and to subdivide these costs into (a) negotiating costs (analogous to those of Demsetz); (b) preemption costs, meaning costs of getting there "firstest with the mostest"; and (c) assertion costs, meaning the costs of simply asserting and supporting a legal claim to exclusivity.

If we consider Christy's paper in the light of this view of costs, we can see that costs of negotiating an international agreement have proved exceedingly high for a number of reasons. There is no clear principle of equity to guide the negotiators even if equitable distribution were an initial agreed objective. Different national wage structures, property systems, and systems of social priorities are difficult to reconcile. The task of working out a unified system of management (including policing) that will achieve the efficiencies that ought to follow internalization is especially obdurate, partly because of the foregoing differences and partly because of the complex and still largely unknown character of oceanic resources. For all these and no doubt many other reasons there has been little progress along the path toward international agreement.

The costs of preemption, on the other hand, are often considerably lower, at least for those with the know-how and wherewithal to follow this course, and certainly less than the gains accruing to the successful preempters. In preemption I mean to include all those activities that are being carried out in order to establish an early and well-entrenched exploitive position—fishing factories, seabed exploration and oil-well drilling (which, even within the acknowledged coastal limits, may drain oil from the international area), research and development on ways to recover the mineral nodules on the ocean floor, and so on. These and other forms of preemption can substitute for legality on the pragmatic principle that, on the high seas, exploitation is nine points of the law.

Nations with valuable coastlines but little preemptive power have taken the initiative in asserting their jurisdiction to greater and greater distances—twelve miles, two hundred miles, and eventually, no doubt, further. The costs of assertion, if not zero, are at least very low. This course is therefore attractive even when the gains, being highly uncertain and elusive, also are low. On the

other hand, policing costs are likely to be high and to inhibit unilateral assertion of jurisdictions that cannot be protected.

The general problem presented in Christy's paper can thus be regarded as one of reducing the costs and increasing the gains of negotiation, while increasing the costs and reducing the gains of preemption and assertion. Since these are to some extent interdependent and time is limited, I shall consider only the first of these objectives.

To reduce the initial costs of negotiation it will be necessary to assure the occurrence and equitable sharing of gains. If the earlier analysis is correct, this would seem to call for (1) finding a key to solving the problem of distribution; (2) circumventing the problem of different socioeconomic systems among nations; and (3) devising an organizational form that will provide effective and efficient management of oceanic resources.

To do this will not be easy, and many alternatives must receive serious attention. The following points are among those that ought to be kept in mind.¹ The first is that the distribution problem is concerned primarily, if not entirely, with the question of how the pure economic rents derived from the ocean's wealth are to be shared. Once this is recognized it may be easier to approach this part of the task with the idea that equity would be truly served only by making every member of the world community a rentier in principle, with a title (through his national state) to the economic surplus due to the natural but by no means indestructible powers of the sea. It should also be recognized that each nation could dispose of its share of sea-rent as it saw fit. This would take care of the problem of equity. Remaining would be (1) the problem of determining "true" opportunity costs, which is a large part of the difficulty associated with differential wage structures, property systems, and values, and (2) the problem of management.

My tentative view is that some version of the corporate form of organization, presumably under international charter, seems likely to offer the most promising approach to dealing with these other two problems. Suppose, for example, that a separate "corporation" were established for each set of exclusive rights, these sets being defined with respect to specified resources in specified regions—such as chasing tuna in certain areas of the Pacific, seals in the Bering Sea, or manganese nodules in defined areas—so that there would be a relatively large number of highly specialized exploitive opportunities and correspondingly specialized corporations. Nations could then bid for the right to operate these corporations and could do so according to their own methods of operating. For example, they could use public corporations or private, pay big wages or low, and so on. The proceeds received from the sale of these

rights would be the initial source of the sea-rents referred to above.

However, for reasons that Vernon believes operated in the cases he discusses, the initial bids are likely to undervalue the rents eventually earned by the successful operators. To determine the excess of "true" value over "true" opportunity costs and to provide ways of capturing it for the world community will be exceedingly difficult. In principle, true value could be established by requiring the producing corporations to auction their output, but this raises many problems. True opportunity costs might be approximated by setting an upper limit on the proportion that value added could bear to value of product (with each corporation free to allocate value added among factors as it wished), but this suffers from rigidity, even with sliding scales, periodic adjustments, and so on. A better arrangement, therefore, might be to maximize the opportunities for renegotiation as a way of eliminating initial and emergent imperfections of valuation. The important role played by opportunities to renegotiate is one of Demsetz' points. Renegotiation is a process which, having been forced *ex post* on the multinational companies, is now leading them, in Vernon's view, to seek various modes of escape which may or may not be successful. It is a process that Albert Hirschman, for reasons that differ from Vernon's, would like to facilitate both *ex post* and *ex ante*. (See *How to Divest in Latin America and Why*, Essays in International Finance, #76, Nov., 1969.) Out of these and other ideas, one would think, a way could be found to permit a periodic recapitalization of economic rents as represented by the current "market" values of various sets of exclusive rights to exploit the seas and to capture the increments of recapitalized value for the benefit of the many rather than the few. A major difficulty, of course, would be to get the socialist nations to accept this as a valid basis of valuation. A second difficulty would be to devise a workable mechanism.

If all this seems like something of a dream world, I suggest reading the monograph by Hirschman. He is telling us that the problems emerging for the twenty-first century call for something better than business as usual. This is the message of Vernon and Christy, too.

One final comment of a general sort. Questions concerning distribution have been fundamental to all three of the papers at this session and have been picked up by the commentators. What emerges clearly is that questions concerning distribution are logically prior to decisions concerning efficiency. Efficiency is the consequence of providing as free a market as possible for the mutually beneficial exchange of "rights"—rights to do, rights to acquire, rights to exclude, rights to impose, rights to consume, and so on. Who has these rights in the first place and how they can acquire them without giving up other rights is a matter of distribution. In short, there must be "possession" before there can be "exchange," efficient or not. Nations and multinational corporations understand this very well, which is why imperialism is still with us.

MARSHALL HALL: Professor Vernon has provided us

¹ After the meetings it was called to my attention that ideas similar to some of those presented here were discussed by Christy and Scott in their pioneering volume on ocean fisheries. See pp. 238-40 in F. T. Christy, Jr., and Anthony Scott, *The Common Wealth in Ocean Fisheries* (Johns Hopkins Press for Resources for the Future, 1965).

TABLE 1
PERCENT CHANGE IN PLANT AND EQUIPMENT EXPENDITURES BY
FOREIGN AFFILIATES OF U.S. CORPORATIONS

	Actual			Projected	
	1965-66	1966-67	1967-68	1968-69	1969-70
<i>By Industries</i>					
Mining and smelting.....	25	17	12	13	2
Petroleum.....	11	19	10	17	15
<i>By Countries</i>					
Latin America.....	2	17	30	26	3
Other LDC's*.....	3	9	16	12	23

* Other LDC's refers to countries other than those in Europe, Latin America, and Canada.

SOURCE: *Survey of Current Business*, Sept., 1969, p. 19.

with some very insightful comments concerning the current and future relationship between the large raw material firm and the less developed countries (LDC's) in which the firms operate. He focuses on what are now the traditional, but still all important, questions of the degree of dependence of the government in the LDC's on the revenues from its raw material sector, the degree of control over the exploitation of its natural resource base by the LDC's, and the sharing of the profit derived from the raw material sector between the LDC and the private firm. His findings and conclusions are also in the traditional vein: the degree of dependence by the LDC's on the revenues derived from their natural resource sector is very great; LDC's are continually trying to increase their share of profit and control; over time they have increased their share of profit substantially, and in some instances the share of profit received by the LDC's cannot increase further without doing harm to both the firm and the LDC. Moreover, the response of the firms and their governments in the more developed countries could be to stress the exploitation of resources inside the boundaries of their own countries by providing tariff protection and encouraging research aimed at making profitable the mining of "low-grade" ore. The LDC's have some alternatives to the response of the more developed countries, but Professor Vernon does not expect these alternatives to be fruitful and he suspects that the control acquired by the LDC's will be illusory with aggregate profits declining.

Without questioning the accuracy of Professor Vernon's conclusions, I would like to suggest that the inevitability of the desire for more control and a greater share of profit by the LDC's has been cast in a framework which stresses the viewpoint of the large raw material firm. It is, I think, this viewpoint which leads to the conclusion that the policies pursued by the LDC's will yield, at best, a Pyrrhic victory.

Consider, for example, the following statement: "Governmental efforts constantly to raise their share of the profits from raw material exploitation can be cited as a near-immutable law of governmental behavior. The problem for governments, however, is not only one of

share but also one of absolute level." I would submit that one could replace "governmental" or "governments" with "private firm" in the above statement and still have a near-immutable law—a law which some would paraphrase as profit maximization. Viewed in this light, the question is simply one of income or profit distribution where the distribution forthcoming will depend to a large extent on the relative bargaining strength of the two groups: the LDC's and the firms. Is 70 percent profit going to the LDC too high? Should the firms, as Professor Vernon seems to be suggesting, always receive the same amount of profit even if their relative share is falling? I think the textbook answer is still very relevant here; i.e., firms, given a choice, will invest in the enterprise which generates the highest expected rate of return taking due account of risk and uncertainty. Seventy percent, 80 percent, 90 percent of profit going to the LDC need not be too high if the rate of return to the firm after the LDC's get their share and after taxes is sufficient to warrant the investment. The proof of the pudding is the eating and despite a possible slow down in investment in 1970 due to political developments in Peru and Chile, the record of the recent past does not indicate a decline in investment in the extractive industries of the LDC's by foreign affiliates of U.S. corporations.

The above is not meant to suggest that at some share distribution (certainly 100 percent of profit going to the LDC's) the private firm will not find investment in the raw material industries of the LDC's unprofitable. It does, however, indicate that one should be wary about implying anything from the fact that the share of profit—more properly economic rent—going to the LDC is 50, 60, or 70 percent. This mode of reasoning is rendered even more difficult when it is recalled that many of the LDC's are integrated firms, and therefore identifying the profit attributable to the "mining" stage is no mean feat.

The response of the raw material firm to the attempts by the LDC's to gain more control and profit that Professor Vernon expects depends, in part, on the willingness of the governments of the more developed coun-

tries to adopt national policies favorable to their extractive firms but not necessarily favorable to their entire society. At the strictly profit maximizing level, if it becomes profitable for the firms to turn inward and increase their domestic investment and research and development in the extractive industries in their own country, then the private firm should do so. In fact, an appropriate question for the LDC's to try to answer is how far can they go in increasing their share of profit without encouraging their foreign firms to change their investment patterns and increase their "at home" investment. The record of new investment in plant and equipment would suggest, however, that that point is still far off. The picture, however, would be quite different if the governments of the more developed countries engage in protectionist policies on behalf of the extractive firms. Putting aside for a moment the question of access to raw materials for defense needs, why should the governments of the more developed countries engage in protectionist policies for the sake of maintaining the profit position of its extractive firms? The market it would seem imposes limits on the ability of the LDC's to appropriate an ever increasing share of the profit from the raw material industries "owned" by foreign private firms. At this juncture in history one would hope that the more developed countries would not engage in policies which foster inefficiency and benefit the large firms in the extractive industries. The sad truth, however, is that Professor Vernon's expectation may well be correct.

Professor Vernon's analysis of the problems stemming from the dependence of the LDC's on revenues derived from the extractive industries is also instructive. To begin with, it lays to rest the notion—which may have been true in previous decades—that the extractive industries are an enclave unrelated to the remainder of the economy of an LDC. This dependency inevitably spawns a desire to control (modify is perhaps a better word) the investment and price or output decisions of what is in many instances the significant industry in the LDC. The accepted tools of control in the more developed countries—monetary and fiscal policy—are, however, of little use in modifying the behavior and

performance of large firms which neither finance their investment nor sell their product in the LDC in which they operate. It would appear that a more direct method of control is necessary. Is it too much to expect that firms accustomed to a fair degree of governmental "control" and exhortation concerning "investment," "price," and "output," in their home country to react rationally rather than emotionally to attempts by the LDC's to evoke the same type of control, albeit by different instruments. Again the investment record of the foreign affiliates of the U.S. firms suggests that after all the shouting they appraise the situation on the basis of profit expectation and despite all the supposed gains in control by the LDC's they continue to increase their investment.

Professor Vernon also comments in passing on the ability of the LDC's to develop state-owned exporting companies that compete with the private extracting firms operating in sheltered, protected markets of the more developed countries. He is not particularly optimistic about the gains although he shies away from making a firm statement one way or another. I agree with Professor Vernon that a unified action by the more developed countries to both protect their indigenous industries and bargain for the "best" deal with the LDC's would present the LDC's with some difficult problems. If, however, that is to be the response of the more developed countries as the LDC's seek to change the profit distribution increasingly in their favor, then the LDC's are in for some even rougher times in the field of international trade and finance, and it is perhaps to their long-run advantage to continue nibbling away at control and profit before the unification takes place.

The above comments are not meant so much to disagree with Professor Vernon's analysis as they are to suggest that if the governments of the more developed countries do not intervene on behalf of their private firms, the outcome of the "inevitable movements" that Professor Vernon observes could benefit the LDC's without hurting the more developed countries. The end result would be income or profit redistribution in favor of the LDC's and away from the large private and multinationally inclined extractive firm.

BASIC DATA FOR POLICY AND PUBLIC DECISIONS: TECHNICAL ASPECTS

DATA, RESEARCH, AND GOVERNMENT

By GUY H. ORCUTT

Urban Institute and Yale University

The paper I have written really consists of nothing more than some views on how the federal government, foundations, and others can best aid the development of more potent, social-problem and policy-oriented research. As you may perceive, this question lies close in my mind to how the federal government might expand and improve its support of data collection efforts which provide the basis for empirical studies of important social phenomenon. In presenting this paper, I start with a deep and abiding belief that economic and social research can make an extremely important—and perhaps vital—contribution to the solution of major social problems. Gains need not be confined to problem resolution though. I also believe that such research can be of vital help in discovering and grasping new and attractive social opportunities.

When I first tried writing this paper—some nine months ago—my thinking was centered on how the federal government could improve the coverage and quality of, access to, and research use of the data it collects. However, the more I tried to put my thoughts on paper, the more I began to wonder if these questions were the right ones.

As I tried to focus my thoughts on the questions, as originally posed, my thoughts went back to the five Federal Reserve committees on data of the mid-1950's, the Inter-University-Ferber effort of the late 1950's, and much of the 1960's, the S.S.R.C. Ruggles Committee, the Dunn Report, and the Kaysen Task Force of this decade. These—and undoubtedly dozens of other studies, to say nothing of innumerable advisory committees—devoted a lot of energy to figuring out and specifying how the federal government might improve the coverage, quality, access to, and research use of data it collects. Each of the above mentioned efforts produced one or more carefully considered and valuable sets of recommendations and suggestions. A few studies have even given rise to significant developments in the actual collection of data. In this connection, it seems reasonable to mention two. I think the Smithies-Federal Reserve Consultant Committee on Survey Research Statistics did assist in making data col-

lected by the Survey Research Center of the University of Michigan more generally available for legitimate research uses. I also believe that the S.S.R.C. Ruggles Committee played an important role in bringing the 1960, 1 in 1,000 sample census tapes into existence.

Both of the above developments have had secondary consequences. No doubt these, along with other beneficial results which do not come to mind at the moment, are enough to justify the dedicated and high-level efforts which have been made. Be that as it may, I have finally come to the conclusion that somehow or other we are missing the boat. The above mentioned efforts resulted in many valuable proposals. Some of these have been implemented and some, like those relating to a computer-oriented national data library, still are of great importance and certainly should be implemented in some form or other. Nevertheless, I now think that a substantially different approach is essential if data needs for social problem and policy-oriented research are to be satisfactorily met.

In medical research, high-energy physics research, agricultural research, defense research, or space research scientists are not in the habit of trying to meet most of their data needs by getting the government to modify its data collection activities and its data access policies. Of course, the federal government has been a major factor behind the amazing development of research in these areas. It does play an enormous role in the data collection activities of and for these sciences, but it certainly does not do this primarily by making data which it has collected for quite different purposes readily available. It helps, stimulates, finances, and provides some direction for research activities in these areas, but it expects researchers to play the primary role in obtaining data as well as in determining which research questions should be posed, what data are needed to answer them, and what strategies should be used to obtain data. Perhaps the soft sciences have something to learn from the hard sciences.

I used the term "soft sciences" for the social problem-oriented sciences quite intentionally, for

the social sciences are in fact very weak sisters with respect to the physical, biological, and medical sciences. This is not to say, of course, that the social sciences have not made significant contributions to the successful handling of some important social problems. Of course they have. Nor is it to say that social scientists are grossly inferior to other scientists. I doubt if they are. But it is correct to say that the cumulation of knowledge about needs, wants, and behavior of individuals, families, organizations, markets, and political and social processes has been hard to see and certainly painfully slow relative to social needs and relative to the cumulation of physical, biological, and medical knowledge. But does this really matter and, if so, can anything be done to improve matters?

Man has been around a long time, judged in terms of his own time scale, and he seems to have managed without much insight from the social sciences. Perhaps he might have managed somewhat better if social science research had flourished much earlier, but this might be hard to demonstrate. In any case, it is not hard to grant that the much greater importance long attached to research and technology, aimed at increasing understanding of and control over nature, may not have been misplaced. Sheer physical survival and well-being are not to be sneezed at. But times have changed and the central problems facing man also have changed. In fact, much of the change is probably due to the very success of physical, biological, and medical research technology.

Population problems are now important and may become much more important. Environmental problems are serious and may become critical. Urban problems are of major concern. War threats now raise questions about man's survival. In part our substantial material success may have brought some problems in the above areas close to the crisis point. But, in addition to aggravating some problems, our material success has made it technically possible to eliminate poverty and promote a greater degree of equality, as well as to deal more effectively and humanely with racial problems and other important social concerns. That some should be grossly in need of food, clothing, medical care, education, and other less essential aspects of the good life must never have been very attractive to those suffering and may never have been very pleasant to think about for those who were not. Nevertheless, the situation seems different and much less generally acceptable when we seem to possess the technical capability for all to live in relative well-being—yet

some of us starve or are ill for lack of effective application of our technical know-how. Surely the verdict must be that it does really matter that our knowledge and understanding about needs, wants, and behavior of individuals, families, organizations, markets, and socioeconomic systems are so limited. We do have innumerable social problems which we would like to solve but whose solution seems to require a degree of understanding about human behavior, organization, and interaction which is simply not available. I do not think we should abandon or even scale down our research in the hard science areas. It certainly is late, however, for our society to be acting as though social problem-oriented research is of relatively negligible concern.

Before addressing ourselves to the question of what the federal government could do to encourage a badly needed development of social problem and policy-oriented research, allow me to indicate what I believe is the major roadblock to development of this type of research. It is the failure of economists and other social scientists to achieve on a wide front even moderately convincing testing of significant hypotheses about individuals and social phenomena. Furthermore, this failure is due largely to the inadequacy of the evidence brought to bear rather than to any great deficiency in our statistical, mathematical, and computer tools. In other words, data and data-related problems are at the heart of the matter.

But to say that data problems are of central importance does not mean that the federal government either already collects the data needed or could be persuaded to collect the data needed. Significant improvements in the collection and use of data by and for government are certainly possible and much to be encouraged. The problem, however, is more basic than that and to some extent involves the whole social science research venture. It relates to a widespread shallowness of social science research when it comes to achieving effective empirical testing of interesting and relevant hypotheses. In my opinion, the shallowness referred to above does not stem primarily from the fact that man and his institutions are more complicated than electrons and atoms, although this is probably true. It results, I think, because collecting of relevant evidence in response to research needs has been largely beyond the capabilities and resources of the individual researcher.

In the hard sciences every student is required to face up to questions which cannot be answered without data and to personally engage in carrying out experiments and in making observations

aimed at acquiring needed data. This process of making hypotheses or of posing questions and then seeking empirical evidence from planned or natural experiments is taught students in grade school and long before the student has finished high school he probably has had quite a bit of laboratory experience. In addition to personal experience, in posing questions requiring data for resolution and in securing the needed data, the student both prior to and during college is taught about important efforts and successes in past application of the experimental approach. By contrast the student in economics typically will never be involved in any activity in which he even participates in securing new data in response to research needs, let alone participate in the designing and carrying out of laboratory or field experiments aimed at generating significant evidence. Students thus trained get their Ph.D.'s and then proceed to train the next generation of economists. If, as professors, economists finally secured some experience in posing hypotheses or questions and in also specifying and actually securing data needed in testing the hypotheses and in resolving the questions, matters would not be so serious. As things stand, however, even at the professional level few individuals ever play a significant role in figuring out what new data are needed for effective testing of hypotheses, how the needed measurements may be made, and in actually making them.

Real and relevant social experimentation is not so hard to think about, but it is hard to see how students could do such experimenting on a routine basis in a social science college laboratory, although even here there are possibilities to be explored. Systematic and careful observation of the enormous number of natural social experiments taking place all the time seems more feasible but also is largely beyond the capabilities and resources of individual researchers. Since empirical evidence is an essential component of hypothesis testing, and, given the difficulty of individual generation and collection of relevant empirical evidence, social science researchers quite naturally and appropriately have sought to make do with data collected by government.

Now there is nothing wrong in using government data for research purposes. Serious efforts have been, are being, and should continue to be made to use the mountains of data collected by government as intelligently and as effectively as possible to throw light on research issues. This is not the time and place to argue the matter. However, I must assert the following. Having thought about this issue from many angles and for more

than a decade, I am now fully persuaded that social sciences will remain stunted until they find far better ways of securing essential empirical evidence than seem possible so long as the prime dependence on government data collection persists. Such data collection of necessity is done for quite other purposes and cannot be effectively controlled by researchers.

Data Need for Basic Research

Basic research aims at discovering the operating or response characteristics of individuals, families, firms, governments, and other key actors in social systems. It proceeds by exploration, by the formulation and testing of hypotheses, by the estimation and testing of relations, and by model building and simulation. It obviously is essential that many of the hypotheses to be tested and the relations to be estimated relate the behavior of decision-making units, such as individuals and firms, to variables which have impinged upon them one or more years earlier. Hypotheses about the influence of pollutants on health, for example, might reasonably involve time lags stretching out over decades. Effective testing of such hypotheses involves several difficulties, but in any case longitudinal data about individuals and their specific environments and habits are called for. In a similar way, theories about the effects of education seek to link the present behavior and success of individuals to previous educational treatment. Clearly, the current or even previous educational treatment given in a census tract in which a number of individuals currently reside will be of rather limited value in testing promising hypotheses. Longitudinal data relating to individuals are needed. The study of migration poses similar data problems. The final conclusion which emerges from the above considerations is that effective study of the response of essential actors or elements of social systems to policy measures requires a data collection approach focused on following individuals, families, and firms, or groups of such decision units over time rather than on following geographic areas over time.

Data Needs for Application of Knowledge Are Different Than Those of Acquisition of Knowledge

The primary differences in data needs have to do with specificity, responsiveness, regularity, longitudinality, depth, and level of aggregation. Data for acquisition of knowledge must be specific to those behavioral or physical units studied, but the units studied may be only a random sample from the population of all such units in the nation.

Data for application of knowledge by urban governments generally must be specific to each and every city or urban area in which application is to be attempted.

Data collection for basic research must be responsive to frequently changing requirements as old hypotheses are set aside and new hypotheses are introduced. Data needs for government also change as the stock of available knowledge grows, but since relatively few hypotheses ultimately prove useful, the requirements of government for responsiveness, with regard to data collection, are substantially less than those flowing from research needs.

Data for use of existing knowledge by government must be timely as well as specific to the geographical area of interest. Thus an emphasis must be given to regularity of collection. This regularity is not nearly so essential for most research uses.

The really difficult aspects of data collection for testing hypotheses have to do with questions of longitudinality and depth. Some significant research requires observations over several years and, in depth, on specific individuals. Since it is dubious if such data collection could be satisfactorily managed by the federal government, let alone the enormous number of relevant urban governments, it is indeed fortunate that use of available knowledge does not pose these particular requirements to the same degree. There are many reasons why use of knowledge poses less severe problems revolving around longitudinality and depth than does research aimed at acquisition of knowledge. Consideration of the acquisition and use of knowledge about the effects of air pollution will bring several of these reasons to light.

First of all, the number of conceivably damaging pollutants is probably enormous relative to the number that cities must actually monitor on a continuing basis for purposes of air pollution control. This is true both because not all pollutants are seriously harmful and because some pollutants nearly always occur together. This knowledge serves to reduce data requirements for operating purposes by telling us what does not have to be measured.

Second, while it may be necessary to have longitudinal data running over several years for specific individuals, to actually discover the harmful effects of any pollutant, use of such knowledge does not pose such requirements. Rather what is needed is current data about the entry of harmful pollutants into the atmosphere. Detail about sources of pollution becomes important and not detail about the previous atmospheres to which

individuals were exposed.

In trying to acquire understanding and knowledge it is frequently essential to work with very disaggregated data. When well-designed experiments are used in seeking to learn how consequences depend on treatment, it is almost always necessary to work with individuals or small groups because cost considerations become dominant. But even with a nonexperimental statistical approach to knowledge it frequently is still desirable to work with very disaggregated data because of considerations having to do with degrees of freedom, use of control groups, avoidance of multicollinearity, and the range of relative variation of variables of interest. In addition, many variables must be included in the analysis as an aid in separating out the true effects of those variables of interest. When it comes to application of knowledge, it will generally be the case that many of these variables which describe individual differences can be ignored, either because they relate only to the past and are thus fixed, or because they do not vary rapidly over time, or because they are distributed randomly over populations and will cancel out at aggregation levels which are of interest in the application of such knowledge by urban government.

The long and short of the matter, then, is that appropriate data collection efforts are well beyond the capacity of individual researchers but cannot be suitably handled by government as we know it. Institutional developments are needed if essential collection of evidence in response to research requirements is to be achieved. Nor should it be thought that collection of needed data for social science research can be bought at a pittance compared to expenditures now being made for collection of evidence by the hard sciences. The rest of this paper presents my present views of the institutional developments which should be promoted.

Proposed National Program for Development and Use of Data Resources for Social Problem and Policy Research¹

Objectives. The overall objective of this program would be to contribute to the solution of social problems and to the discovery and grasping of attractive social opportunities. The more proximate and operational objectives of this program would be the following:

1. Encourage and facilitate research use of

¹ Attention is drawn at this point to the report of the central BASS Committee, "The Behavioral and Social Sciences: Outlook and Needs" (National Academy of Sciences, 1969).

both governmentally and nongovernmentally collected data.

2. Develop and more fully utilize the inductive research potential and problem solving capabilities of college students, undergraduate as well as graduate, along with that of professors and other researchers in both profit and nonprofit institutions.

3. Develop sample survey and other data collection approaches and improve responsiveness to research needs.

4. Promote research use of nonsurvey approaches to observation, measurement and data collection.

5. Promote systematic research use of planned experimentation, both laboratory and field.

6. Promote long-run, basic inductive research and tool building efforts which are focused on finding knowledge of relevance to solution of social problems and on recognition and grasping of social opportunities.

Institutional Proposals. The following institutional developments are proposed for achieving the above objectives:

1. Development of at least one major social problem and policy research institute at or in close proximity to each major university. It is visualized that in general these institutions would be interdisciplinary in nature and would in many cases include physical, biological and medical scientists as well as social scientists. It is expected that they would play a major role in securing research resources, in organizing research opportunities and research facilities, and in providing appropriate mechanisms for selection of research directions, evaluation of research activities and results, and in bringing research results to bear on social problems. Each such institute would be expected to facilitate data collection in direct response to research needs as well as to ensure the local availability of appropriate data library facilities.

2. Development of at least two or three survey research facilities as national, social science research resources. The Survey Research Center at the University of Michigan and NORC at the University of Chicago already have developed a long way in this direction and, of course, would be prime contenders for further encouragement and development as important observatories for the social sciences.

3. Development of perhaps ten to twenty nationally-oriented, major field experimentation and program evaluation facilities. Experience gained by the University of Wisconsin Poverty Institute

and by the Princeton Mathematics Group in planning and executing negative income tax experiments should be invaluable in development of viable and relevant experimentation and evaluation facilities.

4. Development of at least one social science-oriented, national data library and computer software development facility. The S.S.R.C. Ruggles Committee report, the Dunn Report, and the Kay-sen Task Force report all are relevant in this connection. In planning and setting up such a library or libraries, great care should be given to achieving generally acceptable procedures and provisions for guarding against socially unacceptable threats to the privacy of individuals and organizations.

5. Development of several nonprofit, private research institutes for social problems and policy research. The Brookings model, the N.B.E.R. model, and the Urban Institute model should receive close attention in this area.

Support. It is hoped that resources for the proposed program would come from the following sources:

1. Improved research productivity and utilization of existing student and faculty research potential and resources.

2. Partial research redirection of staff effort and funds of the hard sciences.

3. Gifts and bequests.

4. Foundation support.

5. State and local government support.

6. Federal government funding through such federal agencies as the National Science Foundation and the National Health Institutes or through some new social science or social problem-oriented institute or institutes. The level of new funding should start off at a fairly modest figure of less than \$100,000,000 per year. (In other words, at about the level of expected annual operating costs of one projected data collection installation for high-energy physics research.) Nevertheless, the amount of such funding should grow at a reasonable rate, taking into account the time needed to convert and develop human resources, until some kind of appropriate balance is reached between social problem-oriented research and other research whether it is defense-oriented, pure knowledge-oriented, or medically-oriented.

Provisions for Financing, Evaluating, and Guiding. Clearly a great deal of care is needed in working out arrangements for financing, evaluating, and achieving wise and effective guidance. The best I can do at this point is offer a few comments:

1. In general a shift should be made towards institutional and program support and away from individual project support, individual research support, or line budget support.

2. Effective and relevant social problem research requires at least a five- to ten-year horizon. Funding arrangements should take this fact of life into full consideration.

3. Most research planning and implementation should be done within rather than outside of and above research institutes. The primary responsi-

bility for achieving satisfactory individual project evaluation also should fall on the institute within which the research is done. To help in these matters each institute should have an effective governing board which includes some individuals who are heavily involved in social policy making.

4. Evaluation by funding agencies is essential for wise allocation of their resources. Nevertheless, evaluations made should be of the overall direction and operation of the institutes funded rather than of specific projects per se.

MICRODATA, ECONOMIC RESEARCH, AND THE PRODUCTION OF ECONOMIC KNOWLEDGE

By F. THOMAS JUSTER

National Bureau of Economic Research

Introduction

An assessment of the function of microdata in the economic research process—and speculation about its appropriate function in the future—leads to examination of a much broader range of problems. The right question to start with seems to be: What is the function of data in the production of economic knowledge? Putting the question this way suggests other questions: What is the current state of economic knowledge, and how has the character of the data inputs influenced its development? What is the relation between data and knowledge in other scientific disciplines, and if the relation is different than in economics, why? What accounts for the present modes of generating and using data in economic research and are these modes an efficient way to produce knowledge? And, finally, regardless of how the past is evaluated, are there reasons to suppose that the present modes of conducting economic research will be increasingly inefficient in future, and that only a really drastic shift of emphasis—away from concentration on analysis of existing data towards the generation of experimental sets of microdata—will permit much growth in the stock of socially useful knowledge about economic behavior?

The Production of Economic Knowledge

Questions about how to optimize the use of scarce resources to achieve specified objectives have always been of central concern to economists. Analysis of the costs and benefits of alternatives A and B comes as a matter of instinct to all of us, almost irrespective of the subject matter involved. But we have rarely, if ever, looked carefully at the economics of what economists do.

In analyzing the production of economic "knowledge," I prefer to define the term narrowly as the stock of qualitative and quantitative generalizations about economic relationships that are demonstrably "true."¹ Thus knowledge comprises a collection of analytical statements (other things equal, rent control will reduce the supply

of new housing) and empirical statements (the proportion of income saved in the United States is usually between 4 and 8 percent).²

Economic knowledge is produced by professional economists using inputs of their own analytical skills and talents, often in conjunction with inputs of data and statistics. The data inputs are of several sorts.

Basic data can be defined as an empirical observation of a specific phenomenon—perhaps a transaction, perhaps a perception (whether objectively correct or incorrect), perhaps an activity. All basic data are thus microdata, since they must be observations relating to a single economic unit. Observations are defined as basic data that have been generated out of a specific research problem.³ Processed data are basic data that have been adjusted in some way to make them more convenient or conceptually more appropriate for exploration of a specific problem; for the most part, processed data are also aggregates fashioned from microdata.

The production of knowledge will ordinarily require some use of all those inputs, although only the first, analytical skills, seems indispensable. Knowledge can evidently be produced by theorizing about the necessary consequences of a set of relationships, given certain assumptions about objectives. Prior to the 1930's, most of the stock of economic knowledge was produced in exactly this way.⁴ Where data inputs are part of

¹ It is important to distinguish between knowledge and what might be termed "interesting and intuitively plausible relationships," partly because one person's view of what is plausible may be different from another's and partly because the social usefulness of these two kinds of knowledge is probably quite different. For example, demonstrably correct knowledge is likely to have a much greater impact on the formulation of public policy than interesting and intuitively plausible relationships. Moreover, where the latter are used as the basis for policy the social returns are almost as apt to be negative as positive and hence may not average much above zero.

² The distinction between data and observations is discussed extensively in O. Morgenstern, *On the Accuracy of Economic Observations* (Princeton Univ. Press, 1963).

³ As a practical matter, even the purest of pure theorists probably always make some use of empirical information in the research process, even though it may be of a casual and highly impressionistic sort.

⁴ An extensive discussion of different attempts to classify concepts like knowledge and information can be found in Machlup, *The Production and Distribution of Knowledge in the U.S.* (Princeton Univ. Press, 1962).

the process, however, even the selection of one rather than another data input necessarily implies some organizing principle and hence some input of analytical skill.

As a loose generalization, economists produce knowledge increments in one of two ways: (1) theorizing about the logical implications of some type of optimizing behavior; (2) applying a heavy dose of analytical skills to some set of existing data. Knowledge of the first sort is always qualitative and conditional; e.g., higher subway fares will reduce the quantity of subway rides demanded, other things equal.⁶ Knowledge of the second sort is always quantitative and generally associated with specified conditions; e.g., a 5 percent increase in subway fares will eventually reduce the number of riders by 3 percent, provided that bus and taxicab fares remain unchanged. In research that is heavily empirical, a good part of the input of analytical skills may go into adjustment, manipulation, and transformation of either basic or processed data into processed data that has not previously existed.

The Stock of Economic Knowledge

Serious quantitative economic research is a relatively young discipline—a half-century more or less—compared to several centuries for qualitative economic analysis and very much longer for most of the physical sciences. What is a reasonable generalization about the results? If knowledge is defined as pertaining to relationships about behavior, with its implication that behavior can be predicted if its determinants are either known or can themselves be predicted, I suggest that economists possess a very large and often quite useful stock of qualitative knowledge but a remarkably skimpy stock of quantitative knowledge.

If the focus is placed on knowledge useful in the formulation of public policy decisions, as distinct from private decisions of firms and households, the situation is probably even worse. For example, economists “know” that an increase in the minimum wage will tend to reduce the employment of marginally productive members of the labor force and will tend to result in skilled labor and capital being substituted for unskilled labor in the production process. Similarly, it is known that an increase in personal income taxes will reduce the demand for consumption. But neither of these bits of knowledge is particularly helpful in the formulation of public policy, except

insofar as they increase the odds that gross errors will be avoided. Changes in the minimum wage rate will have other consequences, both immediately and in the long run, than a tendency to reduce employment among the marginally productive, and some of these effects will or may be viewed as desirable by some welfare criteria. Similarly, changes in income taxation rates ordinarily comprise one part of a package of policies designed to achieve certain stabilization or equity objectives. Knowing the direction of the effect on, in the one case, unemployment, and in the other, consumption demand, is simply not good enough. What we have to know is the quantitative dynamics of these effects, and we need to know them with a good deal of precision.⁶

Some of these difficulties are probably traceable to the fact that economics has had little success in measuring the influence of social or economic forces that are slow to respond to changes and which represent side effects of policies designed for other purposes.

Environmental pollution, for example, can be viewed as an undesired consequence of growing wealth. Similarly, the attempt to put a floor under the living standards of the urban poor has apparently intensified the social and financial problems of urban areas because of its unintended effect on migration. In short, effective public policies need to be based on general rather than partial equilibrium analysis and on long-rather than short- or intermediate-run dynamics. And even in the simpler world of partial equilibrium analysis and short-run dynamics, the quantitative economic relationships that constitute our best approximations to knowledge are notoriously unreliable and show little evidence of becoming less so. For example, which of our models of investment behavior has predicted the current strength in that sector?

A Knowledge Production Function

To recognize shortcomings in the stock of economic knowledge is neither novel nor especially useful. The existence of serious deficiencies is not

⁶ A qualitative statement is technically identical to a quantitative one that contains very wide limits—if subway rides “decline” they must fall by an amount between one and the present number of rides.

⁶ It obviously matters whether a 10 percent change in the minimum wage will cause a 1 or 100 percent change in unemployment, and we have to specify the time path of the employment effect. And the same is clearly true for changing income tax rates: every undergraduate with a half a year of economics knows that increasing tax rates will reduce the demand for consumption, but recent experience suggests that we lack firm knowledge about the quantitative dynamics of that relationship. In short, for most purposes policy-makers need reliable estimates of empirical parameters, and economists have generally been unable to supply this need.

inconsistent with optimality in resource allocation; economists know better than most that all types of output are subject to benefit-cost constraints, and knowledge is no exception. But perhaps matters are worse than they would have been had research resources within economics been allocated differently.

We can start by examining allocation parameters which describe how research economists spend their time. The alternatives are: (1) theoretical analysis and empirical analysis of existing data; (2) generation of experimental microdata (observations) growing out of a specific research problem; (3) generation of nonexperimental microdata; and (4) generation of processed data, usually but not necessarily aggregates, from microdata. The relevant questions are: first, what are the values of these parameters in the knowledge production function for economics and for other scientific disciplines? Second, to the extent that differences exist among disciplines, is the nature and extent of the difference consistent with a priori and other evidence on efficiency in the research process?

On the first of these questions, hard data are not needed to make reasonable estimates. The proportion of professional skills devoted to the generation of either observations or microdata must be very small in economics (close to zero), while the proportion devoted to generating processed data is much larger than the first two but much smaller than the proportion concerned with analysis.⁷ The great majority of professional economists devote virtually all of their efforts either to the training of future professionals, the study of purely theoretical problems, or the specification and analysis of quantitative problems that use existing, and generally highly processed, data as the sole empirical input.

Costs of Basic and Processed Data in Economics

Some of the forces that influence these choices are clear enough. The generation of economic

⁷ Although most of the professional skill inputs in the basic data function consists of economists employed by the federal government, the great majority of government professionals are concerned with the generation of processed data, not basic data. A few private organizations, for example the N.B.E.R., devote a relatively large fraction of professional resources to the generation or improvement of processed data and a much more modest fraction to generating observations or basic data. Others—for example, the Survey Research Center at Michigan—use a large fraction of resources to generate observations and basic data. And some individuals make similar allocations, mainly on the generation of better processed data. But these are surely atypical allocations of effort for the profession as a whole, although that is probably less true now than it used to be.

data is a classic example of an activity characterized by economics of scale; the economies come in dissemination, not in production. The only reasons for the existence of competitive sets of data seem to be either quality differences or cost differences based on technological lags. Both should cause only temporary interruptions in the attainment of a long-run equilibrium solution where the costs of acquiring existing data are forced close to zero and where a single firm—often the government—produces the product.

The production of new microdata sets is extremely costly. Data from households or business firms are ordinarily obtained by personal contact that always involves the time of the unit being observed and often requires the time of an observer. Both are labor intensive activities, and there has been limited success in increasing efficiency by substituting capital for labor.⁸ Moreover, economic relationships at the micro level are subject to a good deal of nonsystematic variability that is of little intrinsic interest but which is able to obscure the relationship unless relatively large samples of observations are obtained. In effect, disturbances in economic relationships are large relative to the regularities that are of analytical interest, and to insure that regularities can be measured the sample must be big enough so that disturbances are well behaved.⁹ The enormous cost of generating new microdata sets thus means that economists tend to use the best existing sets of data even though they may be seriously deficient for research needs, a tendency that can be offset only if the high costs are balanced by equally high returns.

For the most part, the data inputs into economic research consist of processed rather than basic data, and economics is probably unique among the sciences in the proportion of professional resources that go into the processing and manipulation of basic data. The apparent reason is that much economic data can be fitted into a comprehensive "information system" characterized by a set of logical constraints. In economics, what is bought must be sold, what is produced must be either consumed or added to inventory, what is saved must be invested, and so on.¹⁰

⁸ In some parts of the process—for example, editing the microrecords—mechanical methods have been able to replace labor intensive methods.

⁹ The contrast with the physical sciences, where empirical regularities can often be observed with a sample of one, is quite marked. The noise level in physical experiments is reduced by incurring costs for equipment that essentially serves the function of eliminating disturbances.

¹⁰ The existence of definitional constraints among variables has historically enabled economists to increase socially useful knowledge without making any

The information systems character of much processed economic data is by no means an unmixed blessing. Since an information system is general purpose by definition, the data are unlikely to be optimal for any specific analytical use and will generally have to be modified or adjusted to fit requirements. For example, consumption as measured in the National Income Accounts is not conceptually appropriate for most studies of consumption behavior. It lumps together what basically represent investment decisions like the acquisition of mobile homes, automobiles, and college educations with consumption decisions like the acquisition of food, gasoline, and theater tickets; it blends together real data reflecting observable decisions with imputations reflecting past decisions and estimates reflecting trend extrapolation, and in the process imports a smoothness to consumer behavior that may not exist; and so on.

Second, information systems almost always contain aggregated data, and the system is designed to produce reasonably correct aggregates at the lowest possible cost. The microdata requirements of the system will thus be met, whenever possible, by using existing data generated by some record-keeping requirement.¹¹ Moreover, adjustment of the available microdata has a higher priority than improvements in the quality of the underlying microdata, since accuracy at the aggregate level is not necessarily dependent on accuracy at the micro level because of consistency checks within the system.

An extreme illustration is provided by our savings data. Despite the fact that there are no reasonably accurate microdata, aggregate savings can be estimated with fair accuracy as the difference between income and consumption.¹² And one of our most famous empirical generalizations about savings—that the saving/income ratio is secularly independent of the level of income—is based

use of data inputs at all: it is not essential to measure the money holdings of any economic unit to know that an increase in the supply of money must be held by someone, and that some of the variables which influence the demand for money will therefore be affected.

¹¹ Both the conceptual clarity and the mechanical accuracy of record-keeping data are likely to be obscure unless the data have been generated by trained professionals as part of a specific research design. This is in general unlikely where the data come from a private industry source. See O. Morgenstern, *On the Accuracy of Economic Observations*, 2nd ed. (Princeton, 1963).

¹² This may be an overly optimistic assessment of the accuracy of aggregate savings estimates. Still, it would probably be agreed that estimates of aggregate saving obtained by residual methods are more reliable than estimates obtained from aggregating microdata.

solely on the measurement of net investment in capital assets, which conveniently enough happens to be equal to observed saving.

In sum, cost considerations propel professional economists toward analysis of a relatively small collection of processed numbers that are readily accessible. That this generalization has many exceptions and that it is a better description of the past than the present does not vitiate its essential accuracy. It is broadly correct to describe empirical research in economics as a process that starts with an analytical model, specifies the data necessary to estimate the model, respecifies the model in terms of the best currently available set of processed data, then estimates parameters whose precise quantitative dimensions are heavily dependent on the skill and judgment of the investigator and whose interpretation is thus inevitably ambiguous. It is hardly ever true that economists choose to invest resources in measuring precisely what their analytical models specify.

Data Inputs in the Physical Sciences

An examination of the research process in other disciplines is instructive. In the physical sciences there is virtually no counterpart to the economists' collection of processed data. Empirical research in the physical sciences is based almost entirely on observations generated as an essential part of the research process itself, and a large proportion (probably more than half) of professional skills is devoted to the questions of what observable phenomenon are to be measured and how can the measurement be made. A substantial fraction of the research budgets for sciences like biology, chemistry, physics, and astronomy goes for the procurement of equipment designed to permit extremely precise measurements of observed microphenomenon, and another substantial fraction is spent on equipment whose function is to generate observations. Further, the ability of physical scientists to furnish laboratories with equipment like particle accelerators and cyclotrons designed to produce observable phenomena that can be studied, as well as with devices like electron microscopes, spectroscopes and spectrometers whose function is to permit observation and measurement, is in considerable measure an outgrowth of the research process itself. The power and sophistication of accelerators or spectroscopes are determined largely by what needs to be observed and how accurately the observation needs to be measured, and both are the result of strong interaction between research scientists and equipment builders.

These differences in research methods are well known and hardly require documentation. It

seems plausible that they are also one of the basic causes of differential rates of growth in scientific knowledge. Here, the contrast between developments in space, medicine, and weaponry and the contribution of economic knowledge to the formulation of public policy is rather marked. We can land men on the moon, reduce the incidence of diseases like polio and smallpox to virtually zero, and intercept a potential enemy missile within minutes of its appearance in the atmosphere because we can predict a great many physical consequences with extraordinary precision; but we could not predict the effects of the 1968 tax increase, we could not even agree on the degree of monetary restraint in 1969, let alone predict its quantitative impact, we were unable to foresee the debilitating effects of our present welfare system (and those who did probably also argued that the Social Security system would reduce private saving and thus economic growth), we invest very large sums on a program designed to enrich the educational environment for ghetto youngsters without knowing whether or not the beneficial effects are wholly transitory, and we spend \$50 billion a year on education without knowing whether half or twice that amount is optimal or whether what we do spend is efficiently allocated.

Differences in Research Methodology

While it thus seems reasonably clear that there are marked differences in the growth of socially useful knowledge among scientific disciplines, it is far from clear why these differences exist. The relatively unsatisfactory state of quantitative economic knowledge may be a consequence of the simple fact that economists are concerned with relationships of enormous complexity and have little or no possibility of being able to exercise experimental control. Alternatively, it may result from the relative youth of quantitative economics as a scientific discipline. It might, however, be the result of a sort of methodological hangover that causes a gap between private and social returns to different research strategies.

Experimental Controls in Economic Research

It is undeniably true that the social sciences face limitations in the control of experiments that physical sciences do not generally have to contend with. A controlled experiment requires that a change be imposed on the environment of whatever is being studied and that the consequences of the change be observable and measurable. But economists cannot create a new "Great Depression" just to see how different the result would

have been if monetary or fiscal policy had developed differently than was the case in 1929. Nor can we select students at random to go to four-year colleges while constraining a control sample not to attend four-year colleges in order to measure the net contribution of education to earnings.

This limitation, however, is more apparent than real. Some problems can be explored by analyzing choices under a simulated and synthetic environment—a research design widely used in analysis of managerial decision making and marketing strategy. While it is true that a simulated environment may introduce a bias of important and unknown dimensions, the bias may be the same for alternative simulated changes. If so, and if the real influence of one of two simulated changes is known, the real influence of the other can be estimated from its simulated influence. This design has been used to examine the influence of interest rates on the demand for consumer credit, and a variation has been used to analyze the relation between alternative types of anticipations surveys and actual behavior.¹³

Next, while social scientists cannot make adverse changes in the environment of micro-units in order to observe consequences, it is hard to see the objection to experimental controls involving a favorable change for some units and no change for others. Precisely this kind of experiment is now under way in an OEO Program designed to measure the work-leisure responses of welfare families faced with controlled marginal tax rates.¹⁴ And even the "favorable change or no change" experimental design is not a real limitation. Suppose we do not know if a change will be favorable or not, for example, whether more heavily subsidized charges on urban mass transit systems will expand urban job opportunities, or

¹³ The first illustration is discussed in F. Thomas Juster and R. P. Shay, *Consumer Sensitivity to Finance Rates*, O.P. 88, N.B.E.R., 1964; the second in F. Thomas Juster, *Consumer Buying Intentions and Purchase Probabilities: An Experiment in Survey Design*, O.P. 99, N.B.E.R., 1966.

¹⁴ See Harold W. Watts, "Graduated Work Incentives: An Experiment in Negative Taxation," *A.E.R.*, May, 1969. To achieve the controlled tax rates, OEO is giving away money: sample families who expand earned income are provided with sufficient funds to offset part of the decline in welfare payments. Thus some sample families are receiving more benefits than others, while families not in the experiment are receiving no benefits at all. The experiment is costly. If one wants to reproduce real decisions rather than simulated ones, it is necessary to give away real money rather than simulated money. But there seems no objection in principle to an experiment of this sort.

whether group health insurance plans based on prepayment will provide equivalent health care at lower cost than plans based on fee-for-service. It is not unreasonable to suppose that, in quasi-public firms where optimizing policy is uncertain, a more active searching out of opportunities to influence operating policies could result in the generation of experimental data designed to reduce or eliminate the uncertainties by analyzing the consequences of alternative policies. In short, the limited vision of economists may have made a rigid constraint out of what ought to be no worse than a serious inconvenience.¹⁵

Finally, insistence on absolute experimental control as a prerequisite for scientific investigation is both unrealistic and unnecessary; it is a condition that is often not met even in the "experimental" sciences.¹⁶ In the behavioral sciences, the situation is generally much worse because certain "treatments" cannot be administered at all and because the relation between treatment and "effect" is obscured by uncontrollable "disturbances." But the function of experimental controls is to impound in *ceteris paribus* everything that disturbs the relation between treatment and effect. The best way to achieve this result is to create an environment without disturbances, where the treatment can be administered in varying doses. An alternative is to accept whatever differences in treatment happen to exist in the universe, measure the effects that have been produced by some combination of treatment and disturbances, and then remove the influence of disturbances via statistical procedures.

To do this successfully, however, a number of requirements have to be met. The treatment (say,

change in income) should have a good deal of variation in the observable universe, as should the effect (say, consumption spending). All the disturbances (everything else besides income change that affects consumption) have to be completely specified, and it is essential that disturbances not be too highly correlated with the treatment. And everything has to be measurable and actually measured. The role of theory is critical—to specify the structure of all relevant disturbances as well as the relation between treatment and effect. And as Morgenstern pointed out many years ago, the measurements have to be accurate.

The above set of requirements describes the objectives of quantitative studies in economics. But the objectives are not often met. They cannot be met in any study that uses processed and aggregated data to examine time series relationships, for reasons that have been well known to economists for many years. Nothing is independent of anything else in economic time series, and no one has yet found a way to get around this problem. Additionally, if one believes that there are lots of important disturbances and not just a few, and if one further believes that the relation between treatment and effect is often influenced by the level of the treatment (nonlinearity) and that the influence of one disturbance will often depend on the level of other disturbances and on the level of the treatment (interactions), analysis of time series data is an unlikely prospect for discovering valid empirical relationships.

The shortcomings of existing microdata sets in meeting these objectives, while not so serious as those of time series, are serious enough. Many disturbances are of the sort that cannot be measured directly with existing microdata, and some cannot be measured directly at all because they are unobservable in principle (expected income, permanent income, desired stocks of durables, etc.). Sometimes the influence of treatments or disturbances that are measured will depend on previous disturbances or treatments that were not measured. For example, the influence of high-quality teaching inputs on academic output cannot be detected unless we know the student's ability level to start with. Quality teaching, high test scores, and bright youngsters are quite apt to be found together, and we need to estimate the "value added" by teaching inputs of different qualities, not gross value as reflected by the strong positive association between test scores and teacher quality.

Finally, for virtually all disturbances, treatments, and effects the accuracy of measurement in existing microdata sets is apt to vary from

¹⁵ This general type of experimental design is widely used in private industry research, often for the purpose of determining optimum marketing strategy. For example, the impact of alternative advertising strategies has been studied by varying the advertising dosage for residents of selected geographic areas; then measuring the sales response to get estimates of incremental sales (and profits) per incremental dollar of advertising outlay. Some of these experiments cost enormous sums of money relative to expenditures on basic data generation, and one must assume that they are worth the cost because they continue to be undertaken.

¹⁶ One cannot measure the flow of air in a wind tunnel without changing the flow that is being measured, since the measuring instrument itself represents an intrusion. Similarly, one cannot investigate the effects of genetic composition on cellular growth, which requires identification of the gene responsible for the appearance of mutant strains, without first isolating and then implanting a specific gene in a known complex of other genes—a procedure which in and of itself will necessarily produce changes in the cellular growth process under study.

poor to fair. The economist's equivalent of the electron microscope, the spectroscope, and the observatory seems to be an interviewer with a survey schedule who is instructed to ask: "And what was your family income last year? Under \$3,000? From \$3,000 to \$5,000? Over \$7,500?" Our technology of measurement is, to say the least, primitive.

Differences in Scientific "Age"

As regards the relative youth of quantitative economics as a scientific discipline, it is worth noting that all of the physical sciences with which economics can be adversely compared are, in fact, many centuries older, and their record during comparable stages of development is not especially noteworthy. To illustrate from a scientific discipline which faces precisely the same limitations as economics and other social sciences—inability to control the environment—it took over a thousand years for Copernicus, Kepler, and Galileo to overturn the solar and planetary motion theory of Ptolemy and Hipparchus, and another century or so before Newton was able to put all the pieces together.

Some analogies are worth noting. Development of the elaborate and ingenious Ptolemaic theory bears an uncomfortably close resemblance to quantitative analysis of time series relationships in economics. For example, if a Ptolemaic theorist were faced with an observation that could not be explained by the existing complex set of eccentric circles and epicycles, a new and of necessity perfectly circular epicycle would be invented to explain it. Moderns would of course say that the trouble with Ptolemaic theory was that it could not be tested empirically, since an observation that failed to fit received theory would be duly incorporated into a new and more complex theory. Hence a Ptolemaist never had any "degree of freedom" in the data. Given the amount of analytical ingenuity and computational power focused on the limited number of time series observations available in economics and the strong preference of model builders for linear systems, it is tempting to suggest that the empirical parameters in most time series models bear a much closer resemblance to Ptolemy's eccentric circles and epicycles than to the less esthetically pleasing but simpler (and ultimately more correct) theories of Copernicus and Newton.¹⁷

¹⁷ It might also be noted that the Copernican and Newtonian view of planetary motion was not really accepted by astronomers until it proved capable of successfully predicting events that had not yet taken place. If the adequacy of our econometric models is judged by the same standard, all of us are Ptolemaists

Resource Misallocation in Economic Research

The above analysis suggests that resources used in economic research have been systematically misallocated; more specifically, that the relatively slow growth of socially useful knowledge in economics is a consequence of the fact that research economists have played virtually no role in the generation of economic microdata because of a gap between social and private returns. In essence, the argument is: (1) that the private returns to economic research are not strongly correlated with the discovery of scientifically valid relationships; (2) that many such relationships cannot be isolated without very precise micro-observations specifically designed for the purpose; (3) that the requisite microdata sets are enormously costly relative to existing but inferior alternatives; and (4) that because the costs are much higher and the private returns uncertain and not necessarily higher, economists have strong incentives to use existing and inferior data inputs in research even if the social returns are zero.

The evidence to support this set of propositions can hardly be called persuasive: very little of it is quantitative and much of it can be explained by quite different hypotheses.

Research Design and Private Returns

In research, choices are influenced by the stream of expected income as well as by abilities, aptitudes, and tastes. In economics, as in other scientific disciplines, expected income is related to the quantity and quality of research output. Some of the criteria by which research quality is judged seem to be common to all the sciences and can be summed up by noting that research with a large amount of good theoretical content tends to be associated both with prestige and above average compensation. One can think of numerous reasons for this: correct theoretical specification is the cornerstone of increases in knowledge, and empirical research that lacks this underpinning is unlikely to produce useful results.¹⁸ Once we leave this common ground, however, marked differences appear. Research output in the physical sciences seems to be evaluated on the basis of the skill

—and Newton, to say nothing of Einstein, has yet to make an appearance.

¹⁸ An interesting class of exceptions should be noted. It is possible to relate treatment to effect without knowing anything about the causal relation between the two. A good deal of useful medical knowledge apparently fits this description; e.g., doctors do not necessarily have to know why penicillin reduces the incidence of pneumonia deaths to know that penicillin is a useful treatment for pneumonia.

and ingenuity with which experiments are designed, as well as the theoretical importance of relationships uncovered by the experiment. Any experiment can and will be replicated, often in order to refine or extend empirical results; hence experimental errors are apt to be discovered fairly quickly. Thus the criterion for determining the quality of research output is that it must contribute to a solid base of tested empirical findings within a well-developed theoretical framework.

In economics, however, while analytical skill and empirical ingenuity also tend to be highly regarded characteristics of research output, there is no yardstick for differentiating output that will eventually come to be regarded as part of the fund of accumulated knowledge from output that, while interesting and ingenious, will eventually come to be viewed as wrong. And the private returns to the researcher are apt to be realized long before it is possible to evaluate the ultimate validity of the research output that determines returns.

Let me put the argument more concretely. Assume that you had a dissertation student writing on the question: "Other things equal, are retail prices higher in ghetto areas than suburban areas?" Further assume that data already available could be used to examine the problem, but that the data were, as usual, seriously deficient because they had, as usual, been obtained originally for quite different purposes—for example, say they consisted of total sales and average prices for classes of product and types of stores. The relevant set of observations are a large sample of specific transactions and dates—one jar of X-brand pickles at 35 cents per jar, sold on July 3 at store A, plus other supplementary data. Further assume that an extra two years would be required to collect and analyze the relevant data, and that additional costs would be incurred to raise the required funds. Given these facts, do you recommend that your dissertation student obtain the relevant observations or use existing data? I suggest that your student's professional reputation and expected income would be unaffected by the choice of research strategy, and hence that you could not in fairness recommend the investment of an extra couple of years in generating observations. Yet the social value of the two research designs is markedly different: one has the potential for providing a solid basis for policy decisions; the other does not.

This general line of analysis suggests that the basic reason for the lack of high-quality microdata in economics is the absence of any strong demand for it on the part of professional

economists.¹⁹ While the application of increasingly sophisticated analysis to a limited set of data may have been a reasonable strategy several decades ago, it seems quite unreasonable now—if for no other reason than the fact that the laws of variable proportions and diminishing returns are as applicable to the production of knowledge as to the production of more mundane types of output.

Microdata and Growth in Economic Knowledge

Is there any evidence to suggest that experimentally-oriented microdata inputs will make a big enough contribution to the growth of economic knowledge to justify the costs of obtaining them?

First, all scientific disciplines in which rapid expansion of knowledge has taken place are characterized by the ability to make extremely precise measurements of micropheomenon generated in an experimentally controlled environment. Moreover, the experience of the other sciences is that experimentally generated measurements have a strong feedback on theoretical development.²⁰

Second, while the results of building economic models with good microdata are unknown and uncertain, the results of building models with existing data have been thoroughly explored in a number of areas. The results can fairly be described

¹⁹ It is often argued that relevant but costly microdata sets have not been obtained in economics mainly because of limitations on resources imposed by the unwillingness of foundations and other grant organizations to fund costly projects of this sort. While there is doubtless some substance to the argument, its importance may be exaggerated. The evidence from the physical sciences can be interpreted as suggesting that research resources expand to meet the specifications of research designs, and that costly but essential research designs are not necessarily more difficult to fund than other designs. In short, research resources are responsive to the specification of professionals conducting the research, and the traditional view of an appropriate budget for research in economics is mainly a consequence of what economists themselves have said is necessary. What seems to be fairly common, however, is that great difficulty is experienced in finding resources for the processing and exploitation of existing microdata sets that seem able to provide valuable insights into problems even though they do not qualify as "experimentally generated" observations.

²⁰ Cases in point are Tycho Brahe's astronomical observations, from which Kepler was able to infer that the earth's orbit could be described as elliptical, and current explorations in physics designed to determine whether the shape of the electrical charge in elementary particles like electrons and neutrons is absolutely spherical or only approximately so, which if true would have "profound implication for elementary particles and nuclear structure theory." (See National Science Foundation, *Annual Report*, 1968.)

as unsatisfactory, and there is no compelling reason to suppose that the application of more refined estimation techniques and more fully developed theory to the same set of basic data will result in much improvement.

A few examples will illustrate the point. Large-scale models of aggregate economic activity in the U.S. have been constructed and tested over a period of years. The current models, while more sophisticated in both theoretical content and methods of statistical estimation than those developed a decade or so ago, still do not appear to provide appreciably better short-term predictions than purely autoregressive models with no economic content. The substantive models would not be superior at all if model builders were precluded from making subjective *ad hoc* adjustments to the forecasts produced by the model itself.²¹

Next, despite an immensely large investment of some of our most skilled professional resources, our understanding of the consumption-income relation is not very satisfactory if ability to predict is the test. What is probably the most widely accepted theoretical model is inherently unable to explain the most variable part of consumer outlays (net investment in durable goods and housing), and as yet contains no procedure that is demonstrably superior to any number of other procedures for putting empirical content into its two principal explanatory factors (permanent and transitory income). Quantitative models of aggregate consumption expenditures do not provide appreciably better forecasts than those of a decade ago, and there are still serious disagreements at the theoretical level which cannot be resolved with existing data. In my judgment, one really good set of micro-observations extending over several years would have resolved these questions by now, and probably saved several million dollars worth of analytical resources in the process.

Third, it is possible to specify sets of microdata that are probably capable of being obtained and of resolving issues that have not been (and possibly cannot be) resolved in any other way. This seems to be true, for example, with regard to specification of the consumption-income relation, the effects of minimum wage legislation on unemployment, the monetary returns to education, the efficiency of schooling inputs, and the

long-range consequences of income maintenance programs.²²

Finally, even if high-quality experimentally designed microdata prove to be less useful than one would hope in the resolution of analytical problems, significant gains would almost certainly accrue in terms of sharpening our knowledge of both economic aggregates and of microdistributions. For example, even if it still proved impossible to predict consumption behavior on the basis of experimental microdata specifically designed for that purpose, we would at least have as a by-product a rich collection of consumption, income, and savings estimates that would be available as inputs into information systems.

The Technology of Microdata Generation

One of the most striking differences between the social and physical sciences lies in the technology of measuring empirical phenomenon. Extraordinary advances in the power and sophistication of measuring instruments have been made in the physical sciences.²³ In economics, with the exception of sampling procedures, our measurement technology is very little different today than it was two or three decades ago. A few studies have been made of the errors involved in responses to simple factual questions asked by interviewers—

²¹ One of the distinctive characteristics of the microdata sets with the capacity to resolve these issues is that the observations must cover an extended period of time. If consumption behavior is influenced by expected income and if the latter is in turn influenced by the level and structure of actual income during past periods, precise measurements of both income and consumption covering the relevant past are essential inputs into a properly specified model. Since precise measurement of the past may be impossible, the only solution may be to accumulate data on a current basis until enough time has elapsed. And some of these problems, e.g., measuring the efficiency of schooling inputs, may require even more elapsed calendar time. Here, the output that one wants to relate to school inputs is really post-school performance, and it is likely to be a decade or so after the completion of schooling before an adequate measure of performance can be obtained.

²² In large part these advances are a direct consequence of the demands by research scientists for instruments capable of increasingly finer resolution, as noted earlier. Since the point of many physical experiments would be lost in the absence of sufficiently precise measurements, there is strong incentive for physical scientists to involve themselves directly in the process of designing instrumentation. Much of this emphasis doubtlessly results from the fact that physical scientists are accustomed to specify both the process by which experimental phenomena are to be generated and the procedures by which the phenomena are to be observed and measured.

²³ This conclusion may be unfair to the model builders, since the subjective adjustments might themselves be dependent on the insight provided by the model.

our standard measurement technique.²⁴ But there has been no serious attempt to explore the cost and effectiveness of alternative technologies.

Moreover, existing methods typically assume that the objective is to obtain information from each of a preselected random sample, with little regard to the trade-off between sample design and data quality. In the design of household surveys, for example, use is rarely made of the extensive system of tautological relationships that form the basis of our aggregate information systems data. These surveys are not designed to exploit the fact that changes in assets must equal the difference between income and consumption expenditures provided the consumption categories are comprehensive and asset changes are defined to exclude capital gains, presumably because the survey sponsors do not like to ask respondents to resolve apparent inconsistencies. Yet it is hard to think of a framework better designed to reduce measurement errors than the application of a consistency check.²⁵

In a similar vein, sponsors of microdata collections are seldom prepared to compensate behavioral units for the time and annoyance involved in the observation. In the behavioral sciences, after all, it is not possible to measure relevant behavior without the cooperation and assistance of the unit being observed. But the measurement process inevitably involves at least the cost of time and, for some units, annoyance as well. A straightforward agreement with sample respondents to trade accuracy and consistency in the set of variables being measured against compensation for the time and inconvenience involved, might well be a more efficient procedure than the current practice of appealing to the social conscience of respondents to obtain observations whose accuracy is undermined by an unwillingness to examine internal consistency because of its possible effect on the refusal rate.

Finally, experience strongly suggests that a

²⁴ For example, see E. Scott Maynes, "Minimizing Response Errors in Financial Data: The Possibilities," *J. of Amer. Statis. Ass.*, Mar., 1968, pp. 214-27. Morgenstern, *op. cit.*, is concerned mainly with errors of aggregated data.

²⁵ The statement in the text is oversimplified, since some use is made of consistency checks. A common procedure is to use consistency tests to eliminate cases of doubtful reliability. But the acceptable range is very wide, since ± 10 percent is often viewed as satisfactory when comparing direct consumption estimates with estimated income less savings. But a 10 percent error in consumption means a possible error of many times that percentage in savings estimates.

greater allocation of professional skills toward the problems of generating relevant microdata sets should eventually produce completely new measurement techniques that are vastly superior to the modified versions of existing techniques suggested above. It is probably useless even to speculate about potential developments, but they might possibly be tied to generating observations via preprogrammed computerized methods. If Nielsen can observe television preferences without the need for a physical observer, why cannot economists obtain transactions data in the same way? The point is that the learning process is apt to pay handsome and totally unexpected dividends in the form of much better methods than can now be specified.

Concluding Remarks

Let me close by listing a collection of points ranging from omissions to implications. First, some omissions:

1. The particular research areas discussed above are mainly a reflection of my own interests and do not necessarily constitute the most crucial or the most manageable problem areas. For example, less attention has been given to research on the behavior of firms than to research on the behavior of households: the reason is my greater familiarity with the latter.

2. Although my concern has been with the generation of experimental observations, there is abundant evidence that our general purpose information systems are badly in need of greater resources. To cite two recent cases in point: at a critical juncture in the economy when the consequences of policy errors are likely to be more serious than usual, our basic measure of unemployment has behaved so erratically that it is hard to place much credence in it, and our basic measure of consumption has a well-deserved reputation for looking very different after the annual revision than it did while being reported currently. It is difficult enough to formulate economic policy when the key variables are known within tolerable limits; it is next to impossible when the actual course of events over the past several months will be known only sometime next year.

Finally, some implications:

1. The main thrust of this paper has been that the generation of precise sets of micro-observations growing out of the analytical requirements of a specific research problem is likely to provide a much more promising method of expanding economic knowledge in the future than the application of more sophisticated, analytical, and statisti-

cal techniques to existing data. A strong case can be made for this argument if it is true that accurate micro-observations are in fact obtainable at "reasonable" cost. Evidently, this task may prove to be impossible at any cost, or to be possible but so costly as not to be worthwhile.²⁰ There seems no way to tell except to try, and the right way to try is clearly with a pilot study designed to explore feasibility.

2. If the research process needs to be refocused on generating micro-observations, it is probably true that organizational restructuring will be needed also. Large-scale "research firms" may have to replace the individual "research entrepreneur" to some extent.

3. If the argument in this paper is correct and

²⁰ Two points should be noted. First, to the extent that the experimental microdata required for particular research problems need to be obtained from either of our basic types of micro-units, firms and households, considerable economies of scale in research uses ought to be possible. Using households to illustrate the possibilities, there are few important problems in economic research that do not require precise measurement of a common set of variables over an extended period of time. Investigation of the demand for consumption, the demand for money, the response to price changes, the returns to education, the allocation of time in nonmarket activities, the formation of expectations, the demand for household capital goods, and the determinants of marriage and birth rates—all require collection of a large common set of variables plus a smaller collection of variables specifically oriented to each problem. Thus a single sample could be used to examine a wide range of research problems by adding a moderate collection of additional information. Such a research-oriented sample would have to be large enough so that the observational intrusion on any one unit could be kept within manageable proportions. Second, it may be possible to recoup some of the financial costs by charging fees in excess of marginal costs. For example, the enormously costly Project Talent data appear to be available on this basis. There are thorny problems of research structure and organization involved in this area, however, and they cannot be resolved in this paper.

there is a gap between private and social returns to alternative research strategies, it follows that normal market forces cannot by themselves bring about the necessary reallocation of research resources. The point at which effective pressure can be exerted is probably at the level of organizations that finance research—foundations and government. If projects designed to generate the data required to test hypotheses are treated in a kindly way and other projects are not, one would expect to see more research of the first sort and less of the second.

4. The budgetary implications of the research strategy outlined here are formidable. Looking only at research on households, for example, one can make rough guesses about orders of magnitude. Assume we are really talking about a sample more like 100,000 units than 10,000, about an observation period more like ten years than two, and about a cost per annual observation more like \$100 than \$10. That comes out to a nicely rounded total of \$100 million over ten years or \$10 million per year—just a bit less than the total NSF appropriation for research in the social sciences, but still probably less than NSF now appropriates for the generation of experimental data in physics, in biology, in chemistry, and in astronomy. Just to get some additional perspective, the total is about half the probable cost of the 1970 decennial census.

5. The most significant returns from the research strategy outlined here may not begin to appear until a good many years after the investment has begun, primarily because long-run and general equilibrium effects cannot be observed until a good bit of calendar time has elapsed. And significant social returns are unlikely to appear at all unless professional economists involve themselves, as well as their junior associates and graduate assistants, directly in the data-generating process.

A SYSTEM OF INFORMATION CENTERS FOR RESEARCH AND DECISION MAKING

By CHARLES C. HOLT

Urban Institute

It was with some hesitation that I accepted the request to write a paper on data and computer utilization in the context of research and public policy decision making. The topic is almost too big, too vague, and too important to think about constructively.¹ Typically such efforts have been undertaken by prestigious committees with professional staffs and several years to work. Recent reports from several such efforts² have incorporated both general and concrete recommendations to promote the development of fast and flexible systems for data collection, information handling, computing and communications that can hardly be challenged except in terms of cost, and

¹ Many of the computing and data problems considered in this paper are well recognized and sizable largely independent efforts currently are underway at the National Bureau of Standards, General Service Administration, the Department of Housing and Urban Development, Department of Transportation, Department of Defense, Bureau of the Census, Bureau of the Budget, Atomic Energy Commission, National Aeronautics and Space Administration, Department of Health, Education and Welfare, Bureau of Labor Statistics, the Department of Agriculture, Library of Congress and other federal agencies. A survey by EDCOM[16] reports the following counts of operating time-shared computer networks: 51 university networks, 5 library, 26 national, 24 state-municipal, 4 airline, 3 intrabusiness firm, and 3 interbusiness firm. In addition *Telecommunications* of Aug., 1969, indicates that there now are 132 commercial time-sharing services. Rather than even try to deal explicitly with this tremendous proliferation of divergent developments, we will in this paper attempt to construct a conceptual and organizational framework that is commensurate with the size and complexity of these problems.

² For committee studies that bear on data, information, computing and communication in the social and other sciences see [8] [24] and [27]. For a very comprehensive study of similar areas in natural science and technology see [9] and psychology see [14]. Governmental data and computing are considered in [22] and [17]. This effort was triggered by the American Economic Association and the Social Science Research Council. Individual study of software, libraries and data libraries for the social sciences are given in [2] [3] [4] [16] [19] [23] and [28]. A conference on research and education in information science is reported in [1]. For a general overview of government and social science see [22]. These references should give the interested reader a lead into the still small but rapidly growing literature on scientific information, computation, and communication.

they merit the careful study and full support of our profession. The forthcoming report for the Economics Committee of B.A.S.S. study [24] will contribute further specific recommendations when it comes out.

Since the relevance of both data and computing arise from their functions in the research and decision-making process, it has seemed useful to start by examining these processes in order to clarify the functional objectives. Next we consider in this paper some of the critical problems that need to be overcome. Then we turn to some approaches to their solutions and recommendations on both the institutional and technical levels.

When computer utilization by researchers and decision-makers is carefully examined, it quickly becomes apparent that data processing and analysis is only the beginning. Computers also have a vital role to play in information storage, search, and communication—an area long the province of the library. An examination of the potential complementarities involved in these fields leads me to the conclusion that in the future the distinction between computing centers and libraries will be largely irrelevant.³ Hence my topic is of even broader scope than appeared at the outset.

In addition, the consideration of public policy decision-making needs inevitably involves interrelationships between the various bodies and agencies of government and the multiple levels of government, but also the interdisciplinary dimensions of relevant knowledge, data, and technique.

In spite of the formidable committee competition, a one-man effort may be able to contribute something in clarity, coherence, or incisiveness. In any case, the reader should be warned that I speak as an economic researcher, not as a professional specialist in information technology.

³ The extent to which the production, printing, indexing, and storage of documents is likely to be based on computer based machine readable text is discussed in [20]. There currently is a very active ferment among some of the professionals in library science, computing, and information technology who are struggling to identify, conceptualize, and organize the research and teaching in the new inclusive field of information science. See the proceedings of a recent conference [1].

I. Information, Knowledge, and Policy Decisions

Before attempting to analyze the problems associated with data, computing, and information and proposing solutions, it will aid in the formulation of performance criteria if we examine the larger system that makes use of these tools. Figure 1 sketches an idealized overview of this system which is composed of an interconnected sequence of functions (shown in rectangular boxes). This figure can be interpreted as a crude sketch of how our national information-decision system actually works after a fashion but perhaps it reflects more accurately the functional components of its performance that need to be improved through more effective institutionalization:

1. *Government Operations.* These involve in rich interaction legislation, programs, resources, organizational structures, administrative control, constituency pressures, and evaluation.

2. *Legislative and Administrative Decision Making.* These involve program alternatives, resources, legal constraints, organizational constraints, estimation of impacts on constituencies, and conflicting interest groups backed up by power to influence votes.

3. *Programmatic Research, Planning, and Evaluation.* These involve the application of knowledge about the state of the world, policy goals, program response relationships, planning analyses, and performance analyses.

4. *Problem-oriented Research.* This involves the study of the state of the world, objectives, responses to actions, and methods for planning and evaluation.

5. *Basic Research.* This involves the study of hypotheses about relationships, and basic tools for research and decision making.

Although the diagram is neat, the actual system emphatically is not, so each of these functions needs to be interpreted broadly and flexibly.

The functions are interconnected in two ways (shown by circles). Government operations encounter unmet needs and problems with ongoing programs that pose policy issues requiring decisions. Decision making generates the need for new program or legislative alternatives; i.e., policy problems requiring program design. In turn, program design throws up program problems that require problem-oriented research to determine what courses of action are more effective. This often generates a need for better basic knowledge and methods.

The stimulation of problems moves left in the diagram and the resulting product in the form of

tested knowledge, programs, and actions moves right. Thus each stage usually has two inputs: problems from the next stage right, and knowledge, alternatives, or actions from the next stage left. Also each stage usually has two outputs in the form of problems that are passed left and results that are passed right. The two-way flows between the functions constitute a vital cascaded chain of communication.

All of these functions involve for their performance similar services in the areas of data collecting, storing, and analysis as well as the storage and retrieval of written documents. These are shown schematically as being rendered by information service centers.

It is clear that various parts of the research, decision, and operations process involve (1) searching files for information about reports on operations, laws, library documents, computer programs, and data, (2) extracting the desired information from storage, and (3) displaying or processing it. Completed work typically generates documents that are then entered into the information system and are communicated to the parties affected.

The fact that several functions may be concerned with the same subject matter and analyses means that the same data, computer programs, and documents may have multiple uses. Hence the need for intercommunication between the information service centers. However, the needs of each function also will require specialized programs, documents, etc. Although the demands on these service centers differ, they have enough similarities that they can be supported to a considerable extent by common efforts devoted to research, standards, and a directory of information about sources of data, programs, and documents.

Although the above scheme has been presented in terms of governmental decisions, it is sufficiently general to apply with some modifications to private corporate research, decisions, and operations.

Indeed the contributions of information and research to decision making and operations that are the concern of this paper are not unique to government, so that much of what follows applies to private business and nonprofit institutions as well. However, as necessary as information is for good decisions, in itself it is not sufficient and should not be overemphasized. Clearly the values, institutions, and the will to face problems all strongly affect the quality of the outcome.

With the help of this framework which stresses information, research, and decisions we

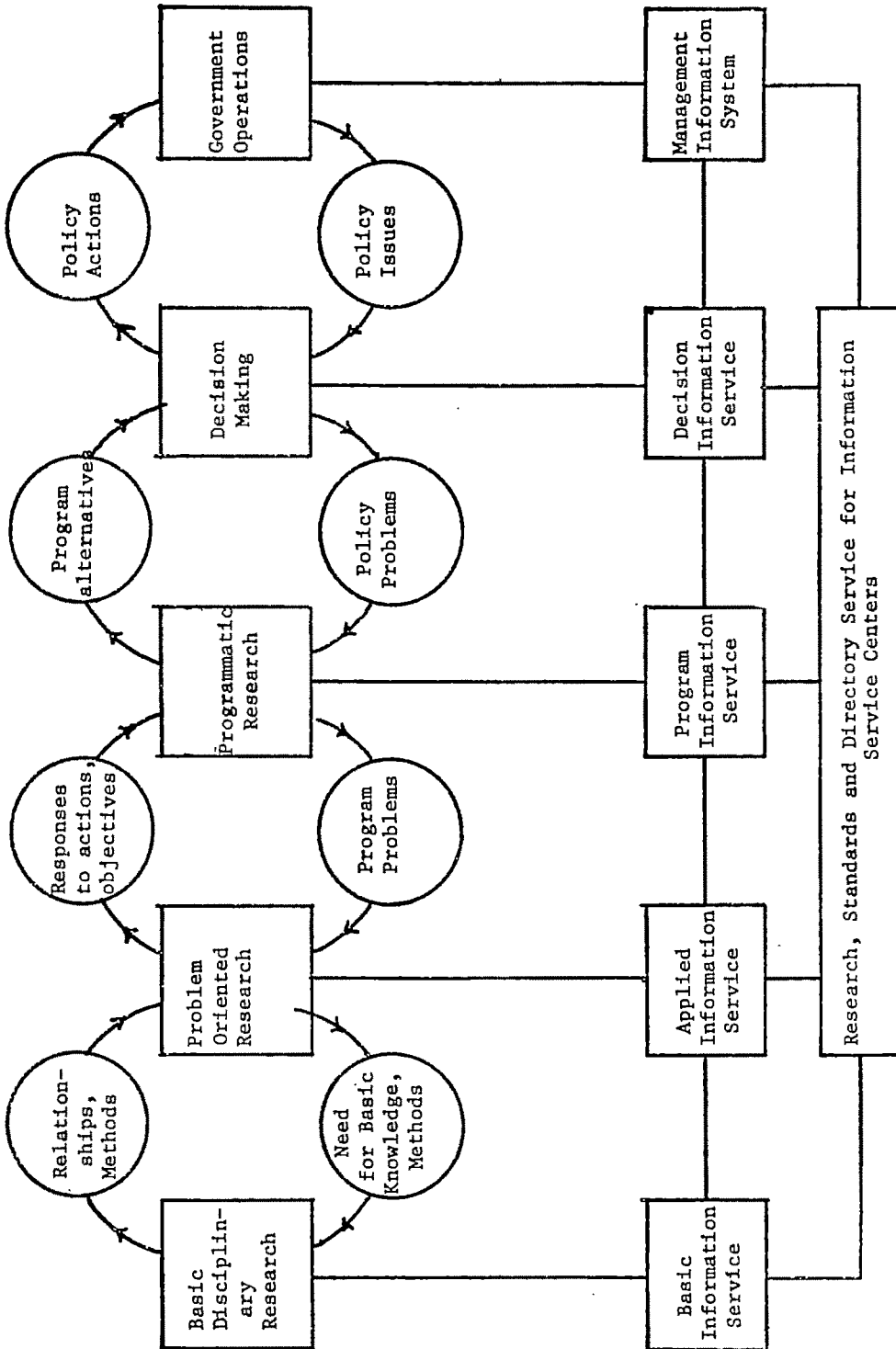


FIGURE 1
NATIONAL SYSTEM FOR GENERATING INFORMATION, KNOWLEDGE AND POLICY DECISIONS

should be able to spot problems and propose solutions in order to further develop and increase the performance potential of our government in helping to solve our complex local, national, and international problems.

II. *Current Problems in the Knowledge System*

Although the system sketched in Figure 1 now exists in a rudimentary form, it is far from being fully effective when measured against what might reasonably be achieved, although not in all instances quickly. The following appear to be problems that merit urgent attention:

1. In the areas of governmental activity for which social science knowledge is relevant, problem-oriented research and programmatic research are relatively weak. There tends to be a serious, largely unfilled gap between basic research and the policy problems that the country faces.

2. In addition there is a severe problem in the communication of research findings and problems between researchers, and governmental administrators and elected officials.

The reason for making these two points⁴ here is that both can constitute such serious gaps in the chain of communication and research as to render sterile the solution of the other problems to be discussed next.

3. Serious incompatibilities between computers, programming languages, and communication codes now seriously inhibit the free and effective flow of programs and data between and among government instrumentalities and researchers.

4. The power of the computer to retrieve sensitive personal data constitutes such a serious potential threat to the invasion of privacy that the development of data systems for research and administrative purposes should (will?) be seriously delayed until adequate legal, technical, and professional safeguards are provided.

5. The costs of relevant data, computing, pro-

grams, and library services are still so great that information services are far from adequate at every stage of the system.

6. Even though tremendous economies of scale are potentially available in the information services, they are seldom realized. The waste from massive duplication of low quality work is notorious. For example, probably two hundred regression programs have been written, but only a dozen are worth running.

7. In the personnel available to staff computing and information activities there often is a serious gap between job requirements and developed capabilities.

To make some of these problems a bit more graphic and to suggest their scale, consider one particular area. Reference [28] is a very good study which reviews computerized information systems for use in metropolitan planning.

The following quotation⁵ pictures the status of general-purpose software for serving this field:

The past three or four years have seen considerable effort devoted to organizing and consolidating the experience gained by individuals and agencies in data base organization, file management, and analysis into comprehensive, generalized systems. A survey of recently announced data management software developments, however, reveals only few systems suitable for urban planning or analysis applications that are available or might soon become available to planning agencies. The major reason for the dearth of suitable software is that much of the advanced development and experimentation in data base systems is taking place in relatively unique computing and operating system environments.

Roughly 190 information systems are listed which are operated by various commissions and agencies. Case studies are presented for 9 of those. The first, an information system in Alameda County, California, has about 2,000 separate computer programs, 50 system analysts and programmers, 30 computer operators, 60 key punch operators and an annual budget for its Data Processing Center of \$2.3 million.

Similar pictures could be drawn of the status of information and computing efforts in hundreds of fields. In aggregate, the scale of resources going into information systems is very large, but few of the complementarities or economies of scale are being exploited.

The stress here on "problems" is not intended to be critical of or to detract from past accomplishments by researchers, government, information scientists, or manufacturers. Indeed, we have

⁴The weakness in applied and programmatic research seriously reduces vital research inputs, and, since effective communications are largely limited by span of concern to adjacent levels, greatly reduces the transmission of problems and the application of knowledge. Both the quantity and quality of basic research in the social sciences is inadequate to our needs. However, there is a great deal more university research at the basic than at the applied levels. In this sense applied and programmatic research are relatively more urgent. Most government programs now are based on "seat of the pants" analysis. Improving the application link should improve the quality and relevancy of basic research, and hence its legitimate claim for greater support. There is no question that much more basic research also is needed to support applied work on many urgent social problems.

⁵ See [28, p. 111].

come a long way in a very few years. Looking forward, however, we risk falling short of attainable objectives by a lack of imagination, crystallization of objectives, and organization of support. The cost of not solving many of these problems is likely far to exceed the cost of doing so.

Probably researchers, policy analysts, and decision-makers see a darker picture in computing and information than people in administrative organizations, because instead of having fairly fixed routines of data collection, processing, and reporting, the former group typically never do the same analysis twice on a set of data. They are forever encountering new programs, new computing problems, and new information sources. Professionals in the information sciences are, of course, thoroughly familiar with these problems and are working actively on them. However, users may be able to cheer and cajole them on to new heights and, perhaps more importantly, help achieve a more coherent allocation of resources.

Next we consider some approaches to the problems listed above.

III. Approaches to Solutions

Because the first two problems listed above are ancillary to the principal topic of this paper, they will be considered briefly.

Development of Problem-oriented Research and Public Administration Education. The deficiencies in applied research and in program planning probably can be traced to: (1) the university is organized around disciplinary departments and professions whose boundaries constitute barriers to interdisciplinary communication, recognition, and support; (2) the social studies have not yet undergone the specialization and professional split of science from problem solving as has occurred between natural science and engineering; (3) most schools of public administration that could be expected to supply a leadership corps of government administrators typically still are obsolescent in terms of modern social science research; (4) the government has not supported the development of interdisciplinary research institutes sufficiently to enable them in any number to buck the university's structure or make it on their own outside the university; (5) universities typically have been unwilling or unable to support organizationally and financially the development of stable interdisciplinary groups; and (6) government agencies generally have yet to build strong in-house professional staffs with adequate time and resources to perform the work needed on program design, program evaluation, and program planning.

The fact that such a cryptic analysis cannot be backed up adequately by good research serves to illustrate the basic point of our relative weakness in applied and programmatic research.

Standards for Articulated Programming. The system sketched in Figure 1 requires ideally the flexible national transmission of data and programs for multiple uses. For example, research on a problem such as the design of an urban transportation system might necessitate the merging of microdata on housing, employment, and recreational facilities. As another example, data on neighborhood population patterns might be used for planning city bus lines, county roads, and state tax offices.

Suppose we were to propose that any person in the national system should be capable of reading and communicating, but perhaps not understanding, any information in data files in the system, and any computer or communications link should be capable of processing or transmitting any file in the system. To cap it off, suppose that each person carried a two-pound portable electronic computer that (although slow) would be capable of reading and writing on any data file.

"Pure fancy," you say. "I can't even get our computer to read multiple punched cards." And you're quite right, of course, but you're talking about today. I'm talking about 1948 when we had some standardized interfaces: the Roman alphabet, Arabic numbers, English language, desk calculators, telephones, teletypes; and nonhardware brains were standard equipment. Now I don't want to be overly romantic about a system that was slow, costly, redundant, and inaccurate, but then we didn't get the following typical stories: "We can read that file but we'll have to write a little program (you know what that means, two weeks or two months) to do it," or, "That program is written for another computer; it would be easier to start over."

To put these problems in perspective we need to recognize that human communication has had several thousand years in which to develop, but communication between computers goes back only a few years. Computer accomplishments already are fantastic and more promising developments lie just ahead. The potential for man-machine systems in the information field presents a tremendous challenge, but a creative synthesis is needed and soon if we are to attain the performance potentials without great pain, conflict, and unnecessary cost.

Coherent systems can develop on an evolutionary decentralized and efficient basis provided that certain critical interfaces are standardized. For

example, once track width was established, the development of a national railroad system could proceed largely on a piecemeal basis with tolerably good results. The development of a national television system required much more detailed signal standards, but once these were established the largely decentralized production of a broadcasting-receiving system was possible and efficient.

When initially poor standards, or worse yet, no standards, are set, investments tend to produce vested interests in parochial standards and these lead to conflicts that are very difficult both politically and economically to resolve. England's problems with measurement units, currency, left-side driving convention, and electric plugs colorfully illustrate the difficulties. Good standards set early minimize later difficulties.

However, this is easier said than done in a field whose technology is undergoing rapid development. Under these circumstances the argument that immediate standardization would be premature has merit. Fortunately, unlike pregnancy, standardization is a matter of degree and of timing. To be useful, standards do not have to be either complete or permanent. Decisions on standards are complex: technically, economically, and politically, in the broadest sense. The most promising approach is to do research on standards at the very frontier of the evolving technology so that good standards can be set early before invested interests become so great that the necessary compromises become both difficult and costly. Grappling with the standards problem early also makes it possible to identify conflicting requirements while it is still possible through further research and development to find integrative solutions suitable for standardization that reasonably meet the needs of the broadest feasible constituency.

In the computing field tremendous investments already have been made and the rate is accelerating in hardware, software, and data systems that are incompatible, so standards⁶ for this field are

late and hence difficult to attain. However, some hopeful beginnings have been made by the computing profession, and the government is helping by putting its purchasing power behind the new standards. However, we have a long, long way to go, and the research resources going into computer standards are utterly inadequate.

One thing that we badly need is a broad conception of where we want to go in setting standards that will contribute most to the development of a national information capability. To suggest the degree of standardization that might be adopted nationally from the continuum of alternatives, consider the following words whose meanings suggest decreasing degrees of organization and system standardization: harmonious, integrated, coherent, consistent, compatible, articulated, unsystematic, unorganized, disorderly, and chaotic. The present situation of computing standards probably is somewhere between "disorderly" and "chaotic."

Wiesen, Yntema, Forgie, and Stowe [30] have developed a system which they refer to as "coherent programming." I propose here a similar scheme that is more ambitious in that it would apply nationally, but admit a somewhat greater range of alternative standards. The word "articulated" carries at least approximately the meaning that I would propose as a suitable target for standardization: "able to tell one's thoughts so they are understood, having joints." This is interpreted for present purposes to mean "intercommunicating, capable of fitting together at interfaces."

Specifications. The objectives of "articulated programming" would be to create a set of standards for computer hardware, program, and data file interfaces so that a decentralized and flexible nationwide network of computers and data files could develop with the following capabilities:

1. Programming languages, file description languages, and computer hardware could be specialized for different kinds of jobs; i.e., an array of alternative languages should be available.

2. Even though programs are written in different languages, they could call each other and call data files expressed in different languages and operating on different machines.

3. Computers could communicate with each other in a variety of languages by instantaneous or batch media requesting computations, programs or files of information in such a way that messages and instructions could be interpreted unambiguously within the computerized system without requiring supplementary human communication or the imposition of a single master language.

4. Data files would be self-described in a vari-

⁶ The need for standards is well recognized in certain parts of the professions [30] [10], Congress [6], the federal government [5], and business [15]. Progress is being made, but the process is slow, painful, and not adequately supported by research. The National Academies of Science and Engineering have some committees working on the standards problem. In a very interesting study of the economics of the computer industry, attention is drawn to the relatively underdeveloped state of software and the associated commercial institutions compared to those for hardware [26, p. 519]. Unfortunately interface standards are not considered at all in this otherwise impressively comprehensive book.

TABLE 1
STANDARDS FOR ARTICULATED PROGRAMMING

I. Program Standard	II. Data File Standard	III. Message Standard	IV. Call File Standard	V. Call Program Standard
1. Language signal 2. Language name 3. Language signal 4. Program name 5. Program arguments 6. Program	7. Language signal 8. Language name 9. Language signal 10. File name 11. User description of file 12. Computer description of file 13. File	14. Language signal 15. Language name 16. Language signal 17. Message to operating system	7-12 from II	1-5 from I

ety of languages with detail adequate both for the user and the computer.

5. When a computer operating system is given program or data files in languages that it is not able to process, it would recognize this condition, reject the job, and make an appropriate response.

6. The standards would apply both to batch process and time shared computers in order to have flexibility between them.

7. When a file or program is called by a program, a check would be made to be sure that the file descriptors, program arguments, and machine configurations are compatible even though different languages may be involved.

8. Documentation on languages, programs, and information files on both programmer and user levels would be current, accurate, adequate, and intelligible.

9. Although all computers, operating systems, and languages would need to be compatible with the above capabilities, purely internal programs and files would not be required to provide them.

That such a system would be desirable can hardly be debated, but the issues of feasibility and cost certainly can. Leaving the cost issue until later,⁷ I would like to indicate in rough outline how these important objectives might be attained through the establishment of suitable standards. Computer specialists may see even better approaches, but I would like to advance this one with the hope of stimulating the consideration of standards that would constitute the framework within which a comprehensive nationwide computing system could evolve.

The articulated programming approach rests on the premise that a single universal program-

⁷ The costs of our present "system," especially its indirect costs, are large indeed. It would not be surprising if, through the introduction of a suitable system of rolling standards, both costs could be lowered and performance improved.

ming and data description language would not be efficient, so that the best we can hope for is a number of languages and the means for translating between them. If we are clever, the number of languages will be small in number, but the computer system in which the languages are embedded should not be so restricted. This open-ended capability is essential because any single language can be expected to undergo a series of evolutionary developments so that it is in fact a series of different languages. The history of the many "standard" FORTRANS is a case in point. Also quite different languages are needed for numerical calculations, information processing, command, instructions to operating systems, teaching and sophisticated programming—to name but a few important areas. The overhead costs of an all-purpose language are likely to be prohibitively high even if we knew how to design one. Upward compatibility is useful, but it falls far short of compatibility.

However, dealing with several languages may be workable if the computer always knows unambiguously in each instance with what language it is dealing and the problem of translating between languages takes place only at limited and well-defined interfaces. Table 1 sketches a system of standards for accomplishing this.⁸ These standards apply to programs (I), data files (II), and messages to operating systems (III). All programs and data files that are to interface with other programs and data files would need to conform to these standards; purely internal programs and files need not.

⁸ In addition to the standards discussed in this paper, other interface standards will be needed for communications, codes, hardware, tapes, etc. For example, it would be desirable to consider designating certain track, density, etc., standards for tape files to enable them to be used to transfer data and programs freely between computers.

A standard "language identifier" would have to be accepted and interpreted by all operating and hardware systems. The language identifier is composed of three parts: language signal, language name, and language signal, in that order (see parts 1, 2, 3; and 7, 8, 9, and 14, 15, 16 in Table 1). The language identifier is a unique bit-pattern, probably a redundant one with a low probability of accidental occurrence that always means the same thing in all systems. It is a separator that encloses and identifies the name of the language. Each language that is sufficiently well defined, documented, useful, and accepted to be admitted as one of the standard languages⁹ is given a unique bit-pattern name. Since any machine could read the language identifier and from it uniquely identify the language, the computer should, if it has the capability to deal with that language, be able to accept and interpret the following program or file or instruction input that is expressed in terms of that language. If the computer cannot deal with an input in that language, it would reject the job and make an appropriate response.

The program, written in the identified language, has the usual name, arguments, and program body (parts 4-6). The data file is described in the identified language by a header which contains: name, user description, and computer description (parts 10-12) and is followed by the file (part 13) that conforms to the structure described in the header. A message to the operating system (part 17) written in the identified language can give instructions to store files, execute programs, assign hardware, etc. In order for the program to be able to call an external data file in such a way that there is an accurate link between the program and the file, the file call must contain an explicit statement of the form of file that the program is written to process. Thus the file call from the program must contain a description in the standard form IV of that file structure expressed in the language of the program that is making the file call. Similarly when a program calls another program it passes arguments to the new program that are expressed in the language of the calling program but using the standard form V to identify the program being called.

To accommodate machine access to programs and files, the operating system would need to maintain a current directory of files and program names and locations. For large files and small computers, this will require interfacing with a human organization to mount tapes, disk packs or other mass storage devices.

⁹ Although machine languages could be admitted, higher-level languages are probably better candidates.

System Operation. Now we can see how the system operates. When one program calls another program that is written in a different language, the program arguments must be passed through by the operating system but the argument expressions must be interpreted in terms of their respective languages. When a program written in one language calls for a data file described in another one, the operating system would need to be able to interpret the structure of the file call and compare it with the structure of the called file to insure that they are compatible, and, if not, possibly interpose a translator between the program and the file at run time.

Operating systems need not only to be able to recognize the language of inputs, but also to recognize whether the input is to be stored away or its instructions executed. This might be handled by designating certain languages as always carrying an immediate imperative to the operating system or alternatively an additional signal could be added to the standard headers. If an additional standard signal were added comparable to the language signal, it might be used to identify each input as "data, program, or instruction to be carried out immediately by the operating system."

One interpretation of the above proposal is that it defines in some detail from the user viewpoint what is needed by researchers and government. It is the responsibility of computer specialists to offer the best system design for accomplishing it.¹⁰ I am under no illusion that the above proposal offers a simple panacea for our computing problems and it of course understates the difficulties. The above proposal may be most fruitful in the context of what we would do with such a system if we had it. This will be developed shortly. The clarification of our critical needs should speed efforts by ourselves and others to satisfy them.

While it is clear that all sorts of difficulties and inefficiencies of such an articulated program system would need to be overcome, it offers a workable strategy for establishing the minimum of standardization necessary for a national information system. It is quite possible that more standardization would be better, but any less would exclude the eventual possibility of flexible nationwide (and ultimately international) inter-

¹⁰ Cheatham, Fischer, and Jorrand in [17] have proposed a strategy in which the provision of a basic compiler language would enable users to build up and compile their own special purpose languages. This research clearly illustrates the extent to which standardization on the right levels and interfaces may enable us to achieve both flexibility and compatibility across machines.

machine communication of programs and files. Clearly the smaller the number of languages and the more compatible their structures, the smaller the number of translations required and the easier they would be to perform. Consequently such a set of articulated standards should contribute to the gradual development of a set of good standard programming languages. The overhead costs for the language identifiers and the other standards here proposed would, of course, require careful evaluation by computer professionals.

The role of the language identifier in the articulated programming system might be illustrated by an analogy. It would be possible for a verbal communication system to change languages between each sentence, provided that an appropriate signal were given at each change: English, German, or French, etc., and that the listener had the capacity to process each of the languages.

Already it is commonplace to incorporate machine language subroutines into FORTRAN programs. Also some interfacing capability has been achieved between FORTRAN, PL/1, and COBOL. What is proposed here is the establishment of interface standards that will enable this capability to be generalized and to be handled automatically by computer operating systems, at least on all "larger" computers.

Environment for Development. The importance of the multilingual translation capability would depend on the particular user group served by the information center. Some centers would need to keep many translators and compilers in high-speed memory while at the other extreme they would be stored for slow access on library tapes, etc.

It would be the rare center that at one time or another did not find the translation capability important in saving program or data efforts.

Tactics have been developed and can be extended for mass-producing translators between languages. For example, this can be done by writing the compiler for a language in the same language (i.e., itself) and then working out a recursive bootstrapping procedure for expressing the compiler in another language that runs on the target machine, either its own machine language or a standard language for which a compiler is available.

Although the resulting machine language program will be slower than optimal, such an approach makes it possible to obtain powerful tested programs in a routine machine intensive way that for many applications will be far superior to starting over with new efforts.

In short, once we have a consistent framework

within which to develop, programs, translation programs, and data files can be produced on a decentralized basis and then widely and cheaply disseminated.

Currently the conversion of programs and files to new hardware configurations involves a good deal of *ad hoc* adaptation by programmers to cope with different word sizes, storage media, communications links, machine instructions, etc. Although it would be complex, computers can be programmed to do a good deal of this on a routine basis and languages could be devised for transmitting the necessary information between machines. Some kind of "machine declaration" needs to accompany programs and data files, and probably should be incorporated in the "message to the operating system." Under this system, programmers increasingly would spend their time designing powerful general systems and dealing with the exceptional machine linkage problems that the general program either could not handle at all or handled inefficiently.

One of the reasons that multiple languages are needed in the articulated programming system is that a hierarchy of functions must be performed and somewhat different language structures are efficient for the different functional levels.

This proposal should not be interpreted as an argument for establishing standards at a very low level. However, we will need to raise our sights considerably even to begin thinking in terms of standards for a national system.

Many nationwide computer-data networks are being established that for their particular purposes require much more rigid standards¹¹ than have been proposed here. A number of academic disciplines have ambitious plans in this direction with NSF support but unhappily the proposed systems are mutually incompatible.¹² For discussion of some of these systems and plans see [4, pp. 16 and 19].

¹¹ The most advanced network system now under development is that being supported by ARPA (Advanced Research Projects Agency of the Department of Defense). To link twenty-odd university and laboratory computers together to share programs and data files, a special interface computer is being used between each computer and the communication network. Other developmental networks are requiring identical computers as the means of achieving the standards necessary for intercommunication.

¹² The diversity of this uncoordinated activity is given by listing some of the groups that are building information systems. King [19] discusses work on chemistry, physics, math, biology, engineering, biomedicine, agriculture, and education. The extensive work on urban information systems was mentioned earlier [28].

The proposed capability that would enable a program or data set to be used by hundreds or even thousands of information centers in spite of having different hardware could drastically improve the quality level of programming, documentation, debugging, and maintenance that would be both feasible and profitable.

The parallel to book publishing is appropriate: the first copy is exorbitantly costly, but the marginal cost of additional copies is extremely low. Such potential economies of scale argue powerfully against current institutions that encourage, even dictate, duplicated efforts. Both copyright protection with billing for services and socially supported production with free distribution should be effectively institutionalized.

How property rights in computer programs can be effectively protected poses very difficult technical and legal issues whose resolution urgently requires research. Instructions to operating systems probably should identify jobs and arrange billing for services rendered.

Articulated programming should contribute to effective decentralization of effort between computer manufacturers and software houses, and increase the effectiveness of competition in both fields.

Protection of Privacy and Restriction of Access. Streamlining the process of gaining computerized access to information immediately raises the issue of protecting information from undesirable uses. There are two aspects of this problem.

One involves the legitimate concern by an individual or organization with information under its control that it not be disclosed to a new user who might abuse the privilege. To protect the information it often is legitimate for its controller to ask what the new user plans to do with the information. Indeed this is the likely price for obtaining access to much information that is at least partially confidential. Access to computerized data can readily be limited to approved people¹³ through the use of code-word keys, cryptographic coding, etc., built into programs and hardware.

The other problem involves gaining access to data and protecting their confidentiality at the same time.

The congressional hearings at the time of the Kaysen report [27] served to sensitize people to the privacy issue as it related to government data but unhappily the important distinction between

administrative data which arises from law enforcement activities directed at particular individuals and statistical data whose analysis supports conclusions about groups of people was never adequately understood.¹⁴

Many social scientists collect data that could be misused to embarrass people and this might endanger the future collection of data. As economists become increasingly interested in microdata at the firm and family level, our profession will increasingly need to concern itself with standards and procedures to safeguard privacy if it is to be socially responsible and maintain access to data sources. The psychologists are far ahead of us in studying the problems of their profession [14].

The problem of protection of privacy in the context of a computerized data library system comes to this question: how can we protect data on individuals while having flexible access for the analysis of general regularities in behavior? Fortunately both objectives are attainable, if we design necessary safeguards into the system. The Bureau of the Census traditionally has solved this problem by keeping all microdata confidential and disclosing tabulations whose smallest cell never contains less than about a dozen individuals. Then presumably no individual could be singled out. Unfortunately, tabular analysis is both crude and extravagant in parameters. A tabulation on only three variables each of which has ten classes yields one thousand cell counts. This is probably much more information than the analysis of the relation between the three variables requires or can use efficiently. It is this large number of "parameters" that tends to make tabular analysis reveal individual information and hence forces the researcher into the consideration of only a few variables, gross measurement classes, and aggregate measures.

Clearly regression analysis with relatively small numbers of parameters is a far superior way both to analyze relationships and to protect the micro observations. A central body of sensitive data that could be accessed only through a regression program¹⁵ could adequately serve the needs of researchers at remote locations, provided

¹⁴So far as I am aware, the merging of all administrative data into a central administrative data bank with personal identification of the merged data was never proposed by anyone. For a recent analysis of the Federal Statistical Data Center concept see [11].

¹⁵The extraction program and provisions for dummy variables for the regression program would have to be suitably restricted in their designs to prevent the researcher from getting at individual observations.

¹³We badly need to improve the legal safeguards for protection of public and private files to avoid undesirable invasions of privacy. See [29]. In many areas the technological ease of communicating data must be rigidly restricted in the interests of protecting civil liberties.

they could give orders to the regression program and retrieve the regression estimates back quickly through their local terminals. This approach would not require the suppression of important information such as the four gross regional classifications in the one-in-a-thousand Census tape. The articulated programming system should facilitate such remote access to protected data bodies.

The Ruggles [27] and the Kaysen [17] committees recommended the formation of a Federal Statistical Data Center that would merge selected bodies of statistical microdata that were generated by some of the administrative and data collecting agencies of the federal government. People have been bothered by the fact that the merging operation would require individual identification of the persons or firms from which the data were collected, thereby making merged data especially vulnerable to abuse.

Aside from the legal safeguards, the confidentiality of the data could be protected from abuse by the employees of the data center if it were never allowed to have individual identifications. This could be accomplished as follows: The data collecting agency could separate at the point of data collection in the field the identification from the data observation by the assignment of a common serial number to both of them. The paired list of identifications and serial numbers would go directly to the identification center that had no data but maintained a master list of identifications in the form of names, social security numbers, etc., paired with a unique set of master serial numbers. The identifications from the data set are matched against the master list of identifications and a new list is produced of serial numbers paired with the corresponding master serial numbers, the identification having been deleted. The new list is then sent to the data center. After the data have been edited and analyzed by the collecting agency, the data identified by serial number are sent to the data center. It then matches the serial numbers from the identification center list against the serial numbers of the data set and produces a new listing of the data paired with master serial numbers. The new data then are merged with old data in the data center by matching the master serial numbers.

This approach has the great virtue of keeping the identification and the data in separate organizations at every stage of the processing after its initial collection. Breach of confidentiality would require collusion between the identification center and one of the data agencies. The files of the data center could be further protected if necessary by encoding them in a form that could only be interpreted by the associated statistical analy-

sis programs so that individual observations could not be retrieved.

An administrative agency, of course, requires individual identification but its data should be legally limited to that needed to perform its particular function. In the case described above they would transmit their identification list to the identification center and their data to the data center, corresponding items carrying a common serial number.

More involved systems could be designed if needed.¹⁶ The above system undoubtedly would be more secure and uniform than present governmental procedures, but even so I know of no breaches of confidentiality over the many years of service by the Bureau of the Census.

A National System of Information Service Centers. There are important economies of scale and complementarities in pulling together into an information service center [3] the functions shown in Figure 2. Libraries of programs, program documentation, data, data documentation, books, periodicals, and research reports involve many common functions: acquisition, storage, control of circulation, cataloguing, etc. As computers become cheaper, the use of automated file searches for data, books, and programs will become increasingly common. Through the use of computers by research libraries and the usefulness of library services to computer users, these two service fields will tend inevitably to draw together even though they start pretty far apart. The closely related services of data collection, design of experiments, and computer programming also could well be merged into the center to supply a full array of information services. While one cannot be complaisant about the difficulties of creating such an organization and training its staff to work as a team, the efficiency and effectiveness of such centers should merit the effort.

The stress should be on a professional organization of sufficient scale to be effective. The present pattern of duplicative, low-quality,¹⁷ and un-

¹⁶ Charles Schultze, for example, has emphasized that only relatively small samples are required for statistical analysis so that there will be a low probability that data would be included from any particular individual. However, there are some technical problems in merging sample files. See discussion in [29].

¹⁷ A study was made of the numerical accuracy of regression programs when making calculations on colinear economic data. The results were fantastically varied and inaccurate. See [21]. Horseback guesses have been made that as much as 80 percent of published computer results are in error to some degree. Computer programs quickly get so complex and testing efforts usually are so inadequate that many programming errors are found years later, if ever. All

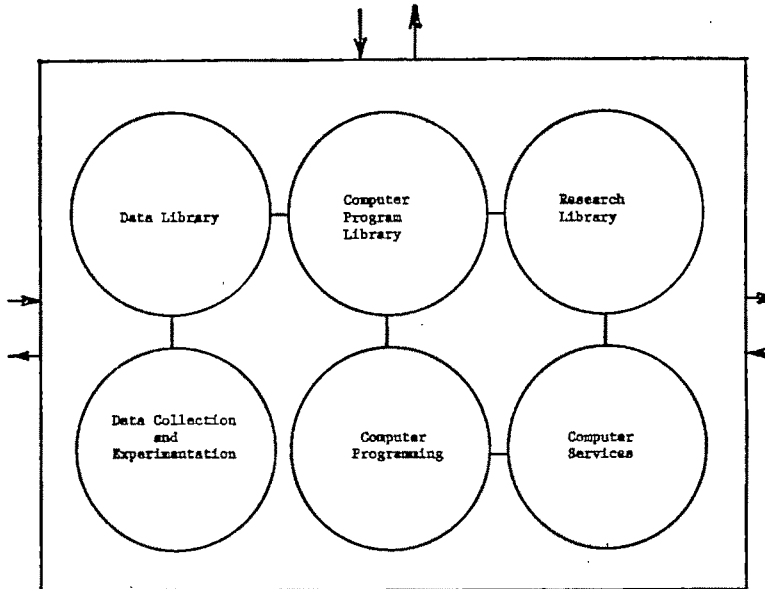


FIGURE 2
AN INFORMATION SERVICE CENTER

derdocumented programming, data files composed of partially labeled boxes of punched cards lying around in corners, and slow, inefficient searchers of library card catalogues or periodical indices leaves a great deal of room for improvement. Some industrial research laboratories are reputed to believe that often it is easier to rerun an experiment than to find a report on previous work in the literature. If programming, data collection, and library indexing are done at high-quality levels and in close communication with their users so that they are also relevant, then in machine readable form they can be reproduced and widely communicated cheaply. If their quality is low, no matter how cheap the cost of transmission, it is indeed better to start over, collect your own data, and write your own programs. As bad as they may be, the errors will be your own.

Quite clearly the programming, data, and library needs differ at different levels of the national system pictured in Figure 1. For example, census data and national income statistics may be fine for detecting problem situations but they may be utterly inadequate for program planning and decision making, or for the study of behavioral relations on the basic or applied levels.

too often if the program "runs," it is assumed to be adequately debugged. However, there are enough examples of high-quality programs that we know what can be done.

What natural science has tried to rely on data collected by the federal government and averaged before release? Economists have for too long, but part of the reason is that they did not have available efficient information service institutions or resources to undertake wide-scale data collection, editing, and analysis.

Because the activities differ at different levels in the knowledge and decision system, the information centers in order to serve them also must be different. Clearly the data, program, and computing needs will differ significantly for information services to physicists, urban researchers, congressmen, agency executives, or city managers. Consequently, what is needed is a national system of information centers¹⁸ dispersed by region, knowledge level, problem area, agency, etc. However, they should be small enough in number that each can be reasonably supported on an efficient scale. One would hope that some careful thought

¹⁸ Congressman William S. Moorhead in Oct., 1968, proposed legislation to establish a Legislative Data Processing Center (H.R. 20422, 90th Cong., 20th sess.) that would tap government data pertinent to congressional operations. The Department of Housing and Urban Development has supported developmental work on urban information centers. Some commercial data networks are now operational and more are planned. However, all of these have been isolated developments with little attention given to interfacing between centers.

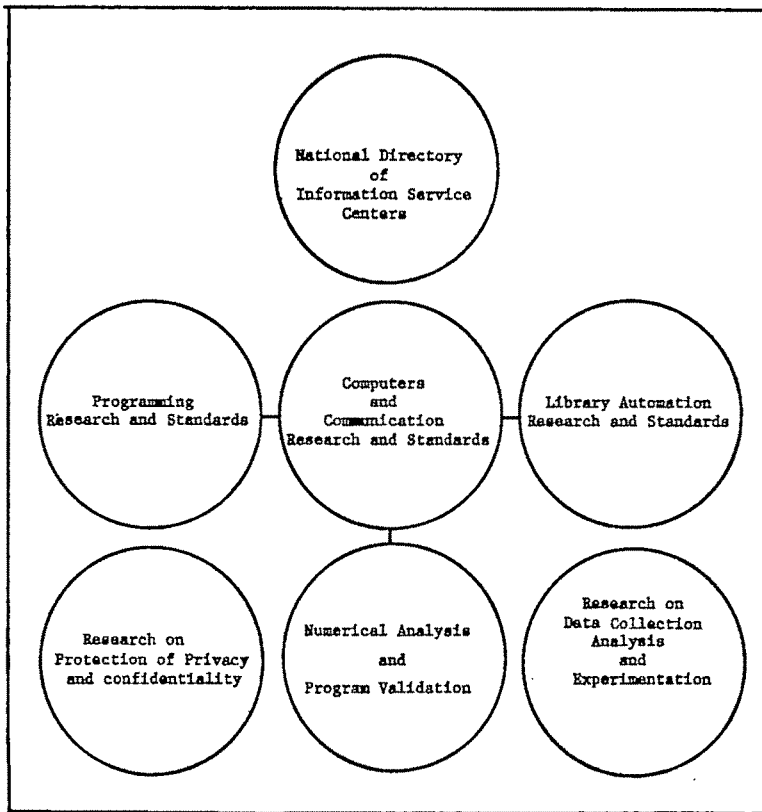


FIGURE 3

RESEARCH, STANDARDS AND DIRECTORY FOR INFORMATION SERVICE CENTERS

would go into planning a pattern that would take account of needs, resources, and complementarities. If nature simply continues to take its course, many of the information centers would end up on route 128.

Some information service centers might be efficiently organized as commercial ventures rather than by the organizations that they serve.

A National Institute for Research on Information Services and Standards. Although the centers will differ, they will also face many of the same problems, and if they are to be of maximum effectiveness, they need to be able to freely communicate data, programs, and know-how. In short, it would be to the national advantage if the dispersed information centers in government, universities, research institutes, and businesses were held together in some kind of a loose intercommunicating system.

To help develop such a system and support it technically, a central activity for research, standards, and directory service is needed. Figure 3

sketches the functions that need to be performed centrally.¹⁹ This effort should be sufficiently large to obtain the economies of scale, and challenging objectives should be set in order to draw in creative talent.

The organization of a central directory activity would make a very important contribution by having in one location information on all the information centers including location, accessibility, and general information on their data, programs, document collections, computer services, data collection services, and programming efforts. This activity could help users find the services they need and help the information centers avoid duplication of effort in data and programming. If programs and data are to be successfully ex-

¹⁹ The production of basic general purpose software for use in the information centers would be one of the important outputs of such an Institute. This aspect of the proposal is very similar to John G. Kemeny's proposal of a National Computer Development Agency [18].

changed between information service centers and their users, careful standards need to be established for the documentation and validation of programs and data.

The federal government has made some scattered moves in this general direction, but the activities relating to hardware and software standards, program development, data, and related research still are seriously fragmented and noncommunicating. However, there are some encouraging signs [12]. The National Academy of Science [8] [9] has been working on some of these problems, emphasizing particularly the need to improve communications between researchers.

In view of the fast changing technology and tremendous potentialities of the information fields, the present system of voluntary standard setting [15] badly needs a broader and deeper research input than has yet been provided. I recommend the establishment as soon as feasible of a federally financed National Institute for Research on Information Services and Standards. An early research project that I would commend to its attention would be the design of a system of national standards that would achieve the goal represented by the foregoing articulated programming proposal—or perhaps go beyond that.

Professionalization of the Information Services and User Involvement. The field of computer science is developing rapidly as a meld of programming, systems analysis, applied math, psychology, linguistics, etc. Its graduates tend to be Ph.D. research-oriented, not service-oriented. The computer science departments and big software research projects often are not very concerned with helping researchers to get their computing done. The assistant professor in the computer science department who would love to program a system for an interesting research project is likely to decide not to because it would not yield a publishable output. Thus some of the most powerful programming talent is diverted from an activity that is in critically short supply and less talented people fill the breach.

We need bachelor's and master's programs in information administration that would merge the library field and the service side of computing.²⁰ These would be primarily teaching programs that would turn out the kind of professional people²¹ that are so badly needed to staff

such activities as information service centers. This program would draw on computer science research, library science, math, statistics, administration, and many other fields. Universities also badly need the people who would be produced by such a program to staff their computing services, but they will have to create for them a better status than the second-class citizenship which typically is now offered. Students can make good programmers but most cannot compete with professional talent. The university will have to make choices between educational or research outputs. For too long, the former has dominated at the serious expense of the latter.

Economic researchers also need to take a more professional attitude toward data, computing, and information services generally. Our professional journals should require as a condition of publication that the author's machine readable data be entered in a data library for ready access by colleagues who wish to test or replicate his results. The author should be able to restrict somewhat the use of data that he collected for a reasonable time until his work with it is complete.

When economists write well-designed and documented programs that are widely used by other researchers or do similar jobs on data collection, these efforts should be recognized as legitimate "publications" even though they are slightly outside our usual activities.²² A good program is a lot harder to write than most economic papers if only for the simple reason that there is a powerful empirical test: it may not run.

Costs. An adequate cost analysis of these proposals is beyond the scope of this paper. Since substantial expenditures already are being made and planned for the collection of information and its manipulation, the cost issue raised by these proposals is one of concept and strategy in directing existing resources as well as of financing entirely new programs.

The benefits from improved information, anal-

²⁰ A wholly new field may be evolving rather than a simple merging [1].

²¹ Currently computer science programs are oriented toward turning out specialists to advance the technology in particular fields of computing. The proposed program would stress the application of existing technology. Consequently its students would

need: (1) a broad background, in contrast to deep, in programming, hardware, and system design, (2) exposure to basic knowledge in an area of application such as accounting, statistical analysis, mathematics, engineering, library science, operations research, or a field of scientific research, and (3) basic knowledge about the organization and administration of man-machine systems. We badly need to get beyond the stage in which a person who has picked up some programming languages is considered adequate to take on any computing job that comes along in any field. The machines are that flexible, but the human staff needs more content knowledge of an area to make the most effective use of the computer.

²² For a parallel, experimental physicists get "credit" for building equipment.

ysis, and ultimately policy decisions need to be weighed against the opportunity costs of diverting high-quality system analysts and information scientists from other activities, of supplying computer, microfilm, and other automated information equipment, of staffing and training the service personnel for information centers (where feasible building on existing organizations), and of responding effectively to the needs of users.

The information centers should be closely tied to their public and private users and many of the costs borne by them. However, a case can be made that information and computer programs are "public goods" in that, once available, the costs of using them are relatively low.²³ Hence economically efficient marginal cost pricing may not cover total costs and public support may be indicated. Similarly the benefits from the research on information services and standards would be so widely diffused that public support is indicated.

In addition to resources, the successful building of such a nationwide system would need talented and creative leadership, and the active cooperation and contributions of many public, private, and professional groups—both producers and users.

IV. *In Conclusion*

In order to see where data and computer utilization relate to policy decisions the interconnections have been traced between basic research, problem-oriented research, program planning, decision making, and government operations.

Problems have been pointed out in the inadequacy of applied and programmatic research, the weakness of education for careers in public administration, communication breakdowns between the research and decision levels, lack of adequate research and standards in computing and information service systems, threats to privacy from computerized data files, low-quality programming, undeveloped service institutions for data and computation, and undeveloped professionalization in the information service field.

These problems have been analyzed and the following recommendations made.

New institutions and programs in applied and programmatic research are needed. The teaching programs in public administration should be modernized and expanded. Standards for articulated programming or an even better system should be adopted to make it possible for a national system of flexible, decentralized, but intercommunicating

information centers to develop. This system should be supported by a federal program of research and standards. New teaching programs are needed to help professionalize the information service fields.

The stress in the whole analysis is on the strong interrelatedness of research, decisions, data, computing, library services, and different disciplines, institutions, and problems. The potentialities of modern computer technology offer a challenging possibility of building a flexible and effective national knowledge-decision system. Such a system should extend to the private sector, but this has not been developed explicitly.

The far-reaching proposals advanced in this paper run the risk of being discredited as "shooting for the moon." We have NASA to thank for converting such challenges from damning indictments into the constructive question, "Which moon do you want to shoot for?" If researchers and public and private decision-makers really want to have good access to data and documents and have computers communicate effectively with each other and with us, I have no doubt that the information scientists and engineers can and will develop the necessary technology if they are given clearly stated objectives, adequate time, and support.

But these issues are much too important to leave to the computer "generals" and manufacturers. Users in government, business, and universities have the critical responsibility for helping to crystalize functional objectives that include the protection of civil liberties and for participating in the process of setting national priorities governing the allocation of support.

However, information services cannot be separated from their research and decision uses. Unless we resolve to improve the institutions and processes for generating knowledge and applying it to the solution of our social and economic problems, a fancy information system can be a nonproductive overhead expenditure. Sad examples are not hard to find: a few urban information systems full of virtually useless data, operating programs that are designed with virtually no knowledge of underlying relationships or clear statement of objectives, and proposals of big computerized operating systems with negligible supporting research on the behavioral relations that are necessary to make them work.

The fact is that we simply do not now have the knowledge base in behavioral social science that is necessary to solve our problems. If we take the actions necessary to meet these research needs, then the information services here pro-

²³ For discussion of the concept of "public goods" see [25].

posed can complement, accelerate, and expedite the process.

But this is not a problem that one profession can solve. It is a national problem that intimately involves government and business as well as researchers, and we need to find some national vehicle to organize attention, leadership, and support for action. The National Academies of Science and Engineering and the Office of Science and Technology are promising possibilities for spearheading national attention to these problems. In view of urgent national problems, it is unfortunate that both still are weak in the social sciences.

To the economist who, not content with considering such global issues, is still trying to get his computer program to run, I can only say: I know how you feel. Don't wait for the millennium. Try prayer.

REFERENCES

1. American Society for Information Science, "Information Science Education: Curriculum Development and Evaluation," Special Interest Group on Education in Information Science, Conference Proceedings of the Curriculum Committee. Jack Belzer, ed., *American Documentation*, Oct., 1969, Vol. 20, No. 4, pp. 326-77.
2. DAVID J. ARMOR, "Developments in Data Analysis Systems for the Social Sciences," Social Science Computing Capabilities Conference, Russell Sage Foundation, New York City, Jan., 1969.
3. RALPH L. BISCO, *On the Development of Social Science Information and Documentation Services (Interim Report)*, Institute of Library Research, Univ. of California, Los Angeles, Dec. 15, 1967.
4. ———, "Role of Computerized Archives for Social Sciences and Humanities," adopted from *Data Bases, Computers, and the Social Sciences*, Bisco, ed. (John Wiley and Sons, forthcoming).
5. HOWARD BROMBERG, "Toward a Standard Standard," *Datamation*, Feb., 1969, p. 22.
6. JACK BROOKS, "The Federal Government and Computer Compatibility," *Datamation*, Feb., 1969, p. 24.
7. T. E. CHEATHAM, JR., A. FISCHER and P. JORRAND, "On the Basis for ELF—An Extensible Language Facility," reprinted from the American Federation of Information Processing Societies, *Conference Proceedings*, Vol. 33.
8. Committee on Information in the Behavioral Sciences, Division of Behavioral Sciences of the National Research Council, *Communication Systems and Resources in the Behavioral Sciences*, National Academy of Sciences, Pub. 1575 (Washington, D.C., 1967).
9. Committee on Scientific and Technical Communication (SATCOM), *Scientific and Technical Communication, A Pressing National Problem and Recommendations for Its Solution*, National Academy of Science, National Academy of Engineering, N.A.S. Pub. 1707 (Washington, D.C., 1969).
10. JOSEPH F. CUNNINGHAM, "The Need for ADP Standards in the Federal Community," *Datamation*, Feb., 1969, p. 26.
11. EDGAR S. DUNN, JR., "The Information Utility and the Idea of the Public Data Bank," *The Information Utility and the Public Choice*, Proceedings of a Conference sponsored by the American Federation of Information Processing Societies in Chicago on Dec. 2 and 3, 1969. (Univ. of Chicago Press and the *Encyclopaedia Britannica*, forthcoming June, 1970.)
12. "Federal Information Processing Standards Notes," *Technical News Bull.*, National Bureau of Standards, Jan., 1970, pp. 16-18. (This article reports on a conference at which nine federal agencies, the National Association of State Information Systems, and hardware and software suppliers participated in discussions of the selection, procurement and management of computer systems.)
13. W. BARKLEY FRITZ, "Computer System Standards for a Large Organization," *Datamation*, Feb., 1969, p. 39.
14. W. D. GARVEY and B. C. GRIFFITH, *Reports of the American Psychological Association's Project on Scientific Information Exchange in Psychology* (American Psychological Asso., 1963-65).
15. ALEXANDER C. GROVE, "Information Processing Standardization: An Evaluation," *Datamation*, Feb., 1969, p. 29.
16. Interuniversity Communications Council (EDUCOM), *Directory of Information Networks and Network Activities (Preliminary Version, First Edition)* (Educational Information Network Program, EIN, Boston, Mass., Mar., 1969, and later revisions).
17. Kaysen Committee, "Report of the Task Force on the Storage of and Access to Government Statistics," Bureau of the Budget (Oct., 1966), reprinted in *The Coordination and Integration of Government Statistical Programs*, Subcommittee on Economic Statistics, Joint Economic Committee, 90th Cong., 1st sess., 1967, p. 196 (alternatively see *American Statistician*, June, 1969).
18. JOHN G. KEMENY, "Large Time-Sharing Networks," *Computers, Communications and the Public Interest*, Martin Greenberger, ed. (Johns Hopkins Press, forthcoming, 1970).
19. HAROLD B. KING, "National Networks," *New Developments in Information Sciences, Conference Proceedings* (North Country Reference and Research Council, Potsdam, N.Y., 1969).
20. FRIEDA B. LIBAW, "A New, Generalized Model for Information-Transfer: A Systems Approach," *American Documentation*, Oct., 1969, p. 381.
21. JAMES W. LONGLEY, "Appraisal of Least Square Programs from the Point of View of the Computer," *J. of the Amer. Statis. Asso.*, Sept. 1967, p. 819.
22. GENE M. LYONS, *The Uneasy Partnership, Social*

Science and the Federal Government in the Twentieth Century (Russell Sage Found., 1969).

23. EDMUND D. MEYERS, JR., "Survey of Social Science Computing Systems," Project IMPRESS (Dartmouth College, June 10, 1969).
24. National Academy of Sciences and the Social Science Research Council, *Behavioral and Social Sciences: Outlook and Needs* (Prentice-Hall, 1970 forthcoming).
25. P. A. SAMUELSON, "The Pure Theory of Public Expenditures," *Rev. of Econ. and Statis.*, Nov., 1954, pp. 387-89.
26. WILLIAM F. SHEARPE, *The Economics of Computers* (Columbia Univ. Press, 1969).
27. Social Science Research Council, "Report of the Committee on the Preservation and Use of Economic Data" (Richard Ruggles, Chairman), reprinted as Appendix 1 of *The Computer and Invasion of Privacy*, a transcript of hearings before the Special Subcommittee on Invasion of Privacy (Hon. Cornelius E. Gallagher, Chairman) of the House of Representatives Committee on Government Operations, July 26-28, 1966 (U.S. Government Printing Office).
28. Systems Development Corporation, *Urban and Regional Information Systems, Support for Planning in Metropolitan Areas*, contract from Department of Housing and Urban Development (U.S. Government Printing Office, M/MP-71, Oct., 1968).
29. ALAN F. WESTIN, "Privacy and Data Banks," discussion by Charles Schultze, *Computers, Communications, and the Public Interest*, Martin Greenberger, ed. (Johns Hopkins Press, forthcoming, 1970).
30. RAYMOND A. WEISEN, DWIGHT B. YNTEMA, JAMES W. FORGIE, and ARTHUR N. STOWE, "Coherent Programming in the Lincoln Reckoner," *Interactive Systems for Experimental Applied Mathematics* (Academic Press, 1968).

DISCUSSION

WILLIAM F. SHARPE: By and large, the authors of these three papers are in agreement. I would like to comment on the general view presented here, and then deal with some questions raised by Professor Holt concerning computation and data services.

The industry that produces economic knowledge is indeed a fascinating one. The product, almost by definition, is conceived to be a public good. Thus it should be (and, by and large, is) supported by government and nonprofit institutions. Much of the support comes from the federal government through research grants and contracts, the subsidization of university computer facilities, and the collection of data. Some support comes (perhaps unintentionally) from state governments via public universities. On the demand side, then, we have something like oligopsony; but the buyers are hardly under the influence of the profit motive.

The firms in this industry—you and I—are small, motivated as much by prestige as by money per se, and fiercely competitive. To produce, they require analytic skills and (maybe) data and computation facilities. The latter inputs, singly or jointly, are held to be subject to economies of scale. The specter of natural monopoly thus looms on the horizon.

Pity the poor economist: an atomistic competitor confronted by oligopsony on one side and real or potential monopoly on the other.

Two major theses have been advanced here. First, we produce too little output; second, our input mix is suboptimal. Professors Orcutt and Holt have made eloquent pleas for more social-problem and policy-oriented research. So has Professor Juster, although he formally admits his inability to argue, as a professional economist, that total output has been too small relative to social goals. On the second thesis there is complete accord. The notion is simply that policy-oriented research should be empirical, and that good empiricism requires the development and use of relatively specialized, expensive, and currently unavailable sets of microdata.

The latter thesis is equivalent to the argument that buyers have chosen the wrong mix of products. We have relatively too much analytic and qualitative economics and relatively too little empirical and quantitative economics. The trick is to change buying habits. Professor Juster would have government and foundations change their preferences among proposals. Professors Orcutt and Holt would directly subsidize the inputs (data and computation) used heavily in the desired type of research. I must admit that I lean toward the former approach, as it seems more likely to lead to a modicum of efficiency in this clearly uncertain and inefficient market.

The key policy problem arising from this discussion concerns government behavior towards data collection, storage, and dissemination. Are the economies of scale so great that any commercial involvement is to be rejected out of hand? I wonder. Professor Holt has indicated that the marginal cost of obtaining usable data has been very high indeed. He would reduce

this via a set of standards and government-supported compatible data centers. And he reminds us of the good old days (1948) when data were just alphabetic characters and arabic numerals and communication took place on telephones and teletypes.

But this is just what the commercial computer industry provides at present. Computers and data files are centralized, with multiplexor devices providing low-cost communication via telephone lines using teletypes and/or similar devices. Files really do seem to be simply a set of readable characters. And users can easily investigate, alter, or list files. Paper tape provides a simple, though slow, means for transferring data and/or programs from one system to another when necessary. And royalty schemes plus file protection allow reasonably full exercise of property rights.

I am more than a little worried by the prospect of a committee-generated set of standards for programs, data bases, and computer systems. Many of Professor Holt's objectives are met by present time-shared systems. Users can consider files simply a set of readable characters. So can programs. To be useful, of course, a file should begin with some explanation of its contents; potential users can then look at the explanation via teletypes. Programs can communicate, although perhaps clumsily, using files. Program A produces data on file X; then Program B reads it. Modern editing systems make it easy to substantially alter files—be they data or programs.

Until recently, even the simplest question (e.g., what is on this tape?) was answered only after several hours or days in a typical computer center. With modern time-shared or remote-job-entry systems, responses can be obtained in seconds or minutes. The sort of standardization proposed by Professor Holt reduces the number of required questions. But it also reduces the number of possible answers. And the cost (time) per question is much smaller than it once was. In this domain I think the forces of competition might prove superior to the deliberations of a committee.

It may be necessary for government or foundations to support the development of data bases and even some specialized languages. But I see no reason to assume that universities, nonprofit institutions, and/or governmental agencies should necessarily provide computer and communications services. There is precious little evidence that such institutions have a comparative advantage in the provision of production-type computer services.

I suggest that we seriously consider the use of commercial services for at least some of our needs. The economies of scale in many areas will allow only limited exploitation by the first (and perhaps only) firm. And the advantages due to competition in data bases, computer hardware, access languages, etc., may more than offset any disadvantages due to monopolistic behavior.

The immediate problem is the education of our customers. We must help show governmental agencies and foundations the importance of sound empirical

research and the development of the data required to support it. Once the market for such research has been developed, we can help educate our potential vendors; then face the make-or-buy decision, utilizing both relevant data and relevant economic principles.

RICHARD RUGGLES: Mr. Juster's central thesis is that the contribution which better quality microdata inputs would make to economic knowledge more than justifies the heavy costs of obtaining them. His arguments and evidence are most impressive. This general view is shared by the other panelists, and I also strongly support this position.

On the other hand, given the present utilization of existing data, I am by no means sure that expensive new collection or elaborate experimentation has as high a payoff as increasing the usability of already existing information. Much of the rationale for creating a National Data System lies in the argument that the vast statistical resources of the federal government are in fact underutilized. If one adds to this the possibility that the design of information obtained in the administrative process could well be altered to be more useful without increasing its administrative cost or burden to the respondent, the potentialities of using data provided by the administrative system are very great indeed.

The present investment in information collected for administrative purposes is not a matter of a few hundred million; but rather runs into the billions of dollars. While not all of this information is of use to the social scientist, the degree of underutilization is such that if it were made available a major resource would open up for economic and social research. It is by no means certain, furthermore, that this would be a costly enterprise, since improvement and rationalization of the federal statistical system might reduce costs, improve the quality of the data, and lessen the burden on the respondent. Until such time as economists have made a serious attempt to improve the federal statistical system, they are in a somewhat weak position to argue for new expensive statistical efforts for which the only function is to provide data for economic research.

In conclusion, I would like to indicate my agreement with Mr. Juster that even if new data collection is not undertaken, organizational restructuring of economic research will be required in order to cope with the massive problems of data processing. Large-scale research firms may have to replace the individual research entrepreneur to some extent, and a team approach which brings together specialists in different disciplines and technologies will be required.

EDWIN KUH: Professor Orcutt's thesis can best be summarized from his own remarks: "... the major roadblock to the development" of vastly improved and more relevant economic and social research is the "failure of economists and other social scientists to achieve on a wide front even moderately convincing testing of significant hypotheses about individual and social phenomena. Furthermore, this failure is due to the inadequacy of the evidence brought to bear rather than to any great deficiency in our statistical, mathe-

matical, and computer tools. In other words, data and data-related problems are at the heart of the matter."

Social science research, he then goes on to argue, is shallow compared to physical science research, although "the shallowness referred to above does not stem primarily from the fact that man and his institutions are more complicated than electrons or atoms, although this is probably true. It results, I think, because collecting of relevant evidence in response to research needs has been largely beyond the capability and resources of the individual researcher."

I feel compelled, partly out of respect for the discussant's role as devil's advocate and partly from genuine conviction, to take issue with this view. These remarks are matters of emphasis—important ones to be sure. However, I most definitely support the basic tenor of Professor Orcutt's proposals. The real crux of the matter is the triviality of physical science subject matter relative to the social sciences, the comparative ease with which many physical science problems bend to formal analysis, the comparative facility with which controlled experiments can be made in the physical sciences, and the absence of value judgments in physical science research. All these differences are intertwined in complex ways that for want of time cannot be entered into here.

All of us want early solutions to the horrendous social problems facing the United States which Professor Orcutt deplors. However, in my view, these solutions more often than not call for the mobilization of political and social will rather more than unraveled social science mysteries. John Gardner has put the matter cogently: "We know our lakes are dying, our rivers growing filthier daily, our atmosphere increasingly polluted. We are aware of racial tensions that could tear the nation apart. We understand that oppressive poverty in the midst of affluence is intolerable. We see that our cities are sliding toward disaster . . . But we are seized by a kind of paralysis of the will. It is like a waking nightmare."

Another point to emphasize—and here Professor Orcutt and I agree—is that the narrow confines of economics and other social science disciplines will balk sensible solutions because of the extraordinary complexity inherent in serious social problems. I do not wish to enter one more sterile plea for "more interdisciplinary research," for the matter is a good deal more subtle than that. Yet we must begin to figure out ways to produce some genuine social scientists, instead of the current crop of excessively specialized practitioners.

Economics has a fair number of successes to record; many important problems within its domain have been solved. But when the domain of inquiry is broadened to include a wide array of social problems, the existing tools help some, but not nearly enough. The most basic problem that I see is analysis, not data. New modes of analysis are needed which combine two contradictory requirements: sufficient complexity to model reality closely enough and enough simplicity to be comprehended. If I knew what this

consisted of, I would be doing it myself. Analysis based on optimization can solve some subset of problems, but the closer we move toward human behavior (as distinct from technological) and welfare problems, the less adequate are these and other available methods of analysis.

Suppose, however, that all the institutional paraphernalia for data generation that Professor Orcutt desires were now in existence. There would be major benefits from research enhanced in this manner, but unless the modes of analysis alluded to are created, we would have more, better data, but no giant new insights into our currently unsolved problems.

Let me illustrate this generality by reference to the Michigan Survey Research Center, a research organization whose quality and contributions have been outstanding in terms of both data collection and their analysis. Much extraordinarily valuable data have been collected, often in response to specific hypotheses that were to be investigated. This is one prototype which Professor Orcutt wishes to have reproduced in larger numbers. Many fine papers have been written based on these data. Knowledge has been advanced in a number of instances, but it was a moderate advance here, another one there, the sort of thing that charac-

terizes most ongoing research. No major breakthroughs of which I am aware have been made, however. The tenuous nature of much of this good quality economic data, the problem of interpretation, and the lack of a sufficiently broad theoretical framework limit the benefits we have obtained from the Michigan Survey Research Center, not the absence of proper data.

Professor Orcutt's second major point is that "social sciences will remain stunted until they find better ways of securing essential empirical evidence than seems possible if prime dependence is placed on data collection by government. Such data collection of necessity is done for quite other purposes and cannot be effectively controlled by researchers." This observation has genuine merit, but should not be overdone. Enormous amounts of valuable statistics are inaccessible and, worse, inconsistent in the present regime of data generation, as Professor Orcutt points out. The research payoff from improved access and codification of existing data would be great. As long as the U.S. government collects vast quantities of pertinent economic data, which it does, strong efforts are needed to systematize and disseminate it to the research community.

PLANNING IN FREE ENTERPRISE ECONOMIES

PLANNING THE FOREIGN SECTOR IN LATIN AMERICA*

By CARLOS F. DÍAZ-ALEJANDRO
Yale University

Recent years have witnessed a reappraisal of planning in Latin America. Regional bodies and scholars have prepared numerous documents reviewing the successes and failures of nearly ten years of intensive planning efforts. On the whole, the balance is disappointing, and there is widespread agreement on the diagnosis. It is asserted that the preparation of a plan document has not been accompanied by the creation of an administrative planning system, capable of implementing the plan and periodically reassessing its targets. The preparation of a plan was in most cases motivated by the 1961 Charter of Punta del Este, which on paper made planning a precondition for receiving U.S. aid. It is charged that long-term plans, whose laborious preparation often had stiff opportunity costs in terms of other public services, have typically been ignored by politicians. Planners could leisurely discuss whether projected growth rates should be 5 or 7 percent, fully confident that the outcome of their debate would not influence any policy decision.

Analyses of the shortcomings of planning have paid great attention to weaknesses in public administration, the delicate relations between planning offices, ministries, public enterprises and the executive, lack of projects and sectorial plans, the absence of consultation with the private sector during plan preparation, weaknesses in the statistical base, and other institutional matters. These are important and relevant considerations, but I doubt that administrative tinkering by itself will improve planning in the mixed societies of Latin America. What has been lacking in these discussions is a clear conception of the priorities and rationale for planning in economies with large market-oriented private sectors. In particular, isolation of key market failures or imperfections which should (and can) be corrected by central planning is a first step seldom taken in Latin American plans. The diagnoses which typically start those plans are little more than a description of the ills of underdevelopment and yield little information on operational policy variables which can be used to remove those ills.

* I am grateful to Mary Downey, James Gough, and Robert Spade for their help in the preparation of this paper, and to my colleagues at the Yale Economic Growth Center for helpful comments.

Take the foreign sector. In most plans, description of past hardships imposed by insufficient and unstable foreign exchange earnings will concentrate on exogenous variables, such as world market trends. This will be followed by projections based on variables whose links with deliberate domestic public policy are at best vague. Time trends, for example, appear with distressing frequency. The question whether the "foreign exchange gap" is due to weaknesses intrinsic in the price mechanism of less developed countries or to misguided domestic policies is left in the air. Even politicians with the best will in the world toward implementing such plans will be left baffled as to how to do so. Should he seek to increase or lower import duties? What should be his exchange rate policy? How far should he go in giving fiscal and credit incentives to private exporters? To these and similar questions the plans give at best Delphic answers as its forecasts are as a rule not made conditional on the levels of those policy variables. So the problem of plan implementation is not just a matter of getting planners and politicians to talk to each other; it is mainly getting planners to worry more about what available policy instruments can affect, and by how much, which economic variables in the context of a mixed economy. If blame is to be allocated for failures in plan implementation in Latin America, the major share of it should be put on planners, not politicians. It is not that the plans are so sophisticated as to be unintelligible to policy-makers (their techniques are typically straightforward); they simply are of little use in answering the questions of economic policy-makers. Politicians reluctant to allow planners to influence policy making even under the best of circumstances (and there are, of course, many of this type) have seldom been confronted with the need to express such reluctance nakedly.

Given the critical role played by the exchange bottleneck and resource misallocation in braking Latin American growth during the last twenty years, this is a sorry state of affairs. It is then surprising that recent discussions give scant attention to coordinating policies aimed at breaking the exchange bottleneck. One such study remarks:

The limitations and obstacles cited have undoubt-

edly detracted from the effectiveness of planning, and can only be overcome by decisions taken in the countries themselves and by the progressive improvement of planning machinery. It should be noted, however, that the implementation of Latin America's development plans has been hampered and is still being hampered by serious obstacles created by the evolution of trade and financial transactions with countries outside the region.

These obstacles are known as "exogenous" factors, *not only because they are largely beyond the country's control* or because they arise from new circumstances which could not have been predicted when the plans were being formulated, but in particular because they alter the economic context in which the plans must be implemented.¹

Planning, like economic integration, means different things to different people, especially in the Latin American context. Some define planning very broadly, as a tool to achieve very long-run economic, social, and political goals, and are relatively uninterested in more modest tinkering to improve the economic system. Others prefer a narrower but more precise definition. Regardless of how widely planning is defined, however, the coordination of domestic public policies toward foreign trade and finance and their compatibility with projected import and export trends should be a central part of planning, especially in countries with a history of foreign exchange difficulties.

Outside Cuba, the Latin American foreign trade and financial sector is largely private, subject to market pressures, which in turn are influenced by domestic public policy. Since at least the 1930's most governments have intervened massively in markets related to foreign trade and have accumulated an impressive array of policy instruments. Planning, therefore, does not start from a *laissez faire* base, and it would appear that one of its first priorities should be to rationalize *ad hoc* policies inherited from the past, seeing whether in fact they are improving or worsening the efficiency of the market mechanism and whether they are helping to achieve other national goals. (The myth of a *laissez faire* initial condition is even less valid in Latin America than in Anglo-Saxon countries, as Hispanic-Latin tradition takes for granted that the state should concern itself with economic regulation.) The rest of this paper will examine summarily some aspects of Latin American foreign trade and capital

movements which could benefit from a planning system which aims at supplementing rather than replacing or ignoring the price system. The foreign sector, of course, is not independent of the rest of the economy and of such matters as income distribution. The theory of the second-best also warns against excessive enthusiasm for taking isolated policy steps in economies replete with distortions, but space limitations forbid a more general discussion of planning.

Planning for Growth and Stability of Exports and Imports

Before discussing foreign trade planning, it will be useful to establish trends in such trade. The often-painted picture of stagnant and unstable Latin American exports and imports is less true for the 1960's than the 1950's. Table 1 presents annual growth rates for Latin American exports and imports, estimated by fitting semilogarithmic trend lines to annual data expressed in current dollars. As a measure of instability around the trend, the standard error of the estimated growth rate is given, in parenthesis. For the sake of comparison, similar estimates for a few non-Latin American countries are included in the table.

For the region as a whole and for most countries, export growth was higher during 1959-68 than during 1948-58, although few countries outside Central America reached during the latter period the export expansion rates registered in the United States and other industrialized countries (excepting the United Kingdom). More remarkable was the decline in export instability, as measured by the standard errors of the growth rates.²

² It is of interest that countries having higher export growth rates do not typically show higher standard errors until very high export growth rates are reached. In fact, as shown in the following tabulation, up to those high growth rates a (very slight) declining tendency in standard errors can be found:

AVERAGE STANDARD ERRORS OF EXPORT GROWTH COEFFICIENTS

	1948-58	1959-68
<i>Annual Export Growth of:</i>		
Less than 3 percent.....	1.56	0.98
3 to 8 percent.....	1.26	0.93
8 to 12 percent.....	1.14	0.90
More than 12 percent....	1.67	1.43

The hypothesis that fast growth (in exports or GNP) is more stable than low growth deserves further empirical study.

¹ United Nations, Economic Commission for Latin America, "Planning in Latin America," *Economic Bulletin for Latin America*, Vol. XII, No. 2, 1968, pp. 14-15 (emphasis added).

PLANNING IN FREE ENTERPRISE ECONOMIES

171

TABLE 1
ANNUAL GROWTH RATES OF LATIN AMERICAN IMPORTS AND EXPORTS, 1948-68
(Percentages)

	1948-58		1959-68	
	Imports	Exports	Imports	Exports
<i>All Latin America...</i>	4.53 (1.00)	3.96 (0.65)	4.22 (0.48)	5.00 (0.24)
Argentina.....	-1.05 (1.88)	-3.39 (1.92)	-0.63 (1.47)	4.87 (1.11)
Bolivia.....	1.89 (1.32)	-4.69 (1.76)	10.32 (0.67)	13.32 (1.44)
Brazil.....	1.19 (2.14)	1.13 (1.32)	2.39 (1.83)	4.38 (0.72)
Chile.....	4.46 (0.88)	4.42 (1.40)	7.22 (1.16)	8.32 (0.94)
Colombia.....	5.86 (2.38)	5.65 (1.85)	2.55 (1.48)	2.15 (0.69)
Costa Rica.....	9.85 (0.70)	6.25 (1.15)	8.78 (0.80)	8.49 (0.69)
Dominican Republic..	9.63 (1.80)	6.08 (1.09)	7.49 (3.32)	-0.06 (1.61)
Ecuador.....	8.85 (1.61)	10.69 (1.62)	7.77 (1.52)	5.22 (0.75)
El Salvador.....	10.86 (0.84)	9.57 (1.12)	9.92 (1.14)	7.98 (0.59)
Guatemala.....	7.56 (0.81)	8.77 (1.20)	8.19 (1.01)	9.24 (1.01)
Haiti.....	2.79 (1.34)	1.22 (2.00)	0.36 (1.36)	1.41 (1.40)
Honduras.....	7.72 (1.10)	1.87 (1.04)	11.69 (0.92)	12.01 (1.02)
Jamaica.....	10.63 (1.23)	13.24 (1.28)	7.72 (0.65)	5.53 (0.91)
Mexico.....	8.19 (1.03)	6.46 (1.00)	6.99 (0.62)	6.05 (0.42)
Nicaragua.....	15.00 (1.43)	14.23 (2.05)	13.68 (1.05)	12.52 (1.42)
Panama.....	4.85 (0.62)	11.38 (1.71)	10.34 (0.45)	14.99 (1.82)
Paraguay.....	2.23 (1.45)	1.10 (0.74)	9.01 (1.24)	7.31 (1.57)

TABLE 1 (*continued*)

	1948-58		1959-68	
	Imports	Exports	Imports	Exports
Peru.....	8.40 (1.30)	6.90 (1.05)	9.67 (1.72)	9.89 (0.96)
Trinidad.....	7.70 (0.89)	9.25 (0.79)	5.57 (1.12)	5.95 (0.60)
Uruguay.....	-0.26 (2.03)	-3.41 (2.14)	-2.78 (1.65)	4.80 (1.68)
Venezuela.....	9.08 (1.37)	8.87 (0.42)	2.58 (1.93)	2.22 (0.21)
<i>Selected Non-Latin American Countries</i>				
Australia.....	2.80 (1.45)	0.90 (0.96)	7.60 (0.81)	6.93 (0.57)
Canada.....	7.62 (0.88)	5.87 (0.87)	8.95 (1.35)	9.45 (0.79)
Denmark.....	6.41 (0.49)	7.60 (0.46)	8.70 (0.62)	7.78 (0.44)
France.....	6.25 (1.00)	8.06 (1.20)	10.78 (0.50)	8.47 (0.41)
United Kingdom....	3.08 (0.93)	3.98 (0.48)	5.34 (0.40)	4.88 (0.23)
United States.....	6.18 (0.89)	5.97 (1.19)	8.62 (1.05)	7.13 (0.43)
Netherlands.....	8.17 (0.85)	10.96 (1.10)	9.35 (0.42)	9.16 (0.29)

SOURCES AND METHOD: Basic data on dollar values, at current prices, of exports and imports obtained from International Monetary Fund, *International Financial Statistics*, several issues. For each period, growth rates were obtained by fitting the following regression to the data:

$$\log Xt = a + bt$$

Where Xt refers to exports or imports, and t to time. The coefficient b will yield the growth rate; its standard error will provide a measure of instability. For Chile, Ecuador, and All Latin America, import data do not include 1968. Data for All Latin America refer only to countries shown in the table. During 1966-68 four countries accounted for 61 percent of Latin American exports: Venezuela (24 percent), Brazil (15 percent), Argentina (12 percent), and Mexico (10 percent).

In only three Latin American countries was the absolute value of that statistic higher during the more recent period. The same trend can be observed in non-Latin American countries suggesting the key importance of factors exogenous to Latin-American economies in yielding greater stability during the 1960's as compared to the 1950's. It appears hard to argue from the export

data that Latin American economy policy during the 1960's faced a more difficult external environment than during the 1950's. Import prices, it may be noted, rose less during the more recent period.

Table 1, however, also suggests that the Latin American foreign trade and payments situation is far from rosy. While during 1948-58 imports

grew faster than exports for the region as a whole, the opposite has been the case during 1959-68. In fact, overall import growth declined between the two periods (although a different picture emerges if Venezuela is left out of the total). In spite of faster export growth and the expansion of net official capital inflows during the 1960's, foreign debts accumulated during earlier years have in many countries yielded heavy interest and amortization outflows, which have limited the capacity to import. Some countries have also devoted large sums to rebuilding exchange reserves. Table 1 also shows that even during 1959-68, export instability in most Latin American countries remained substantially more severe than in the sample of industrialized countries shown. That difference, however, disappears when Latin America is taken as a whole, suggesting one advantage of close Latin American economic cooperation. The table also shows, incidentally, the heterogeneity in export and import performance hidden behind regional averages.

Exports. Given foreign debt obligations and target GNP growth rates of at least 6 percent per annum, the need for most countries to raise export growth rates even above those registered during 1959-68 is fairly clear. What role should planning play in raising the rate of export expansion? Although there are still some who argue that Latin American exports are determined overwhelmingly by exogenous world trends and by the domestic structure of production, not easily changed over the medium run, evidence is accumulating that the observed export rate of expansion is also markedly influenced by domestic policies. The diversity in export trends shown in Table 1 among different Latin American countries and between periods for a given country cannot be solely explained by luck. Peruvian, Mexican, and Central American growth rates compared with those of Argentina, Brazil, and Uruguay for 1948-58 reflect different domestic policies. So does the contrast in export performance of Argentina, Bolivia, and Brazil between 1959-68 and 1948-58. Colombian experience over the last two or three years also indicates that domestic policies do matter for export growth.

If haphazard statism is associated with planning, as often is done, it could be argued that export promotion in Latin America needs less, not more, "planning." Whatever the failure of free market prices may be in reflecting true opportunity costs, the distortions introduced by public policy appear to have been greater in this field. A first step toward a more rational policy will include the establishment of exchange rates which more closely approximate opportunity costs and

give incentives, across the board, for new exports. The variety of goods which enter or could enter foreign trade, as well as the variety of actual and potential foreign markets, strongly suggests the wisdom of decentralizing the search for new export lines. An *ex ante*, centralized search for promising new export lines which would then receive special encouragements is likely to yield fruits slowly.

There is a danger, however, that in the reaction against past policies it will be argued that export policy should consist of little more than a favorable exchange rate. But several well-known reasons indicate the need for considerable public sector intervention in the markets for exportable products. If nothing else, there is the venerable optimum tariff argument which still has empirical relevance for several Latin American traditional exports. That foreign demand price elasticities have often been underestimated in the past by Latin American countries does not prove that they are always infinite. For commodities for which international agreements exist, national planning becomes an unavoidable consequence of international commitments. Indeed, a country facing those international agreements without careful strategy and tactics, developed at the national level, is likely to lose ground in its traditional markets. Other traditional Latin American exports, such as oil and copper, are marketed under conditions which are far from purely competitive, so the preparation of a rational bargaining position by each country (or a group of them) can substantially improve its gains from trade.

Markets for nontraditional exportable goods can be improved by policies going beyond the establishment of the "right" exchange rate. Public action will be required to keep distortions in other sectors of the economy, including import duties and controls, from limiting export growth, at least while those distortions cannot be eliminated directly. Drawbacks, credit and tax concessions can all be used for that purpose. The public sector can help offset lack of credit and information regarding foreign markets that hamper the growth of "infant exporters." A larger share of planning talents could be directed toward designing and quantifying the impact of new public policies of export promotion. In a world economy dominated by powerful trading blocs and nations, which incidentally do not rely just on normal market incentives to promote their exports, the Latin American countries would do well to develop their own export-promoting infrastructure, including the diplomatic front. Because such infrastructure is not likely to be built quickly

enough by the free play of market forces, public planning can play here a useful role.

The classic rationale for planning emphasized lack of future prices for most goods, which could coordinate efficiently today's investment plans. The point remains valid when applied to exchange rates. This consideration can lead either to building large models yielding shadow exchange rates for the foreseeable future, or to a policy of freely fluctuating rates, on the assumption that expectations will somehow be reflected in today's rate. Neither recommendation is likely to be operational in the Latin-American context, either because of lack of data or due to imperfections in the foreign exchange market. Yet the problem of determining the right exchange rate remains a real one, especially for those countries receiving large but unsteady capital inflows, and it deserves the attention of planners.

Planning, in short, need not be the enemy of export growth and foreign trade. Defined as an instrument which corrects market failures and inefficiencies, it can in fact stimulate exports. Furthermore, if planning is also used, as it should be, to offset possible regressive effects of foreign trade policies on income distribution, it will help to neutralize a traditional source of political opposition to export promotion.

Instability of export earnings is one of the oldest topics in discussions regarding the Latin American economy. As shown in Table 1, this instability declined in recent years; in fact, for many countries of the region as well as for those outside, import instability has been greater than that for exports. This fact suggests that much can be done by domestic policies to eliminate excessive fluctuations related to foreign trade. Besides seeking to reduce unexpected shifts in the net supply of foreign exchange, efforts can be made to make it more elastic with respect to policy instruments, so as to facilitate the adjustment to exogenous shifts which will remain in spite of export and market diversification.

Imports. It is by now well known that:

With very few exceptions, the Latin American countries cannot be said to apply a protectionist policy, if by this is to be understood a systematic body of measures deliberately designed to permit and encourage the development of certain industries rationally selected within an overall framework of objectives established under a given economic development policy.³

The maze of *ad hoc* import restrictions and

tariffs accumulated over the last thirty years should be a priority target for planning. Planning offices have made some efforts along these lines, but the results have been meager. Where some tariff reform has been accomplished, it has usually relied little on planning offices. This has been due not just to the fact that tariffs are the province of powerful ministries jealous of their primacy; it is also due to most planners' lack of interest in issues like tariff reform.

Disappointment with the results of import substitution policies followed by many Latin American countries during the last thirty years is widespread among economists of all schools of thought. The lessons derived from this experience, however, are quite varied. One "lesson" often heard is that the experience demonstrates the inability of Latin American entrepreneurs to generate a rapid pace of healthy industrialization. This ignores the biases induced by public policy to which profit maximizing entrepreneurs responded. If effective rates of protection are 300 percent for activities producing luxury consumer durable goods and negative for industries which could export, it is not surprising to find too few manufactured exports and too many consumer durables. Before writing off local entrepreneurs or launching a search for grandiose new strategies to substitute for the decrepit one of import substitution, it may be worthwhile to attempt a rationalization of public policies providing the framework for private activity.

Such rationalization will involve two tasks: one, the determination of an ideal long-run policy of protection, taking into account administrative limitations, and including the creation of mechanisms for periodic tariff reviews. The issue here will not be whether or not to protect, but exactly how much protection can be justified by infant industry considerations and by long-term goals of the development plan. The other could be labeled planning for the transition and would try to minimize the cost of moving from present to ideal policies—a task which could take several years. Close coordination will be required during that transition between exchange rate and import liberalization policies and perhaps also with policies toward foreign finance.

Latin American experience as well as theory indicates that the instruments used to achieve import substitution have a large bearing on its success in promoting efficient activities. It makes a difference whether automobile production is stimulated by uniform tariffs or by prohibitions and whether one or several firms are allowed to operate within the country. Furthermore, there are straightforward and practical techniques for eval-

³ Santiago Macario, "Protectionism and Industrialization in Latin America," *Economic Bulletin for Latin America* (United Nations), Vol. IX, No. 1, Mar., 1964, p. 61.

uating import substitution investment projects (as well as export-promoting ones), such as the "exchange cost criterion" used in Israeli planning,⁴ which could be profitably used in Latin American planning offices. The application of theory and these techniques, incidentally, could satisfy both the planner's interest in sophisticated tools, as well as the policy-maker's search for operational criteria. Techniques for making projections and for insuring the internal consistency of plans, such as input-output, need not be the only tool in the planner's kit.

*Toward Latin American Economic Integration:
How Much Planning?*

Among the many factors hampering the progress of Latin American economic integration one finds a false dilemma between a "planned" integration and automatic regional trade liberalization. It is argued that Latin American trade liberalization should come after, or at least should be subordinated to, the achievement of regional development plans. Given the limitations of planning even at the national level, one may wonder whether this viewpoint is simply a way of postponing integration indefinitely. Most national plans, incidentally, say little on the possible impact of integration on national economies; typically optimistic export projections are justified partly with general references to integration, while import forecasts tend to ignore that possibility.

At the root of that view is a failure to differentiate between trade liberalization and liberalization of factor movements, as well as between sectors where market forces can be expected to yield socially acceptable results and others where this is an unlikely outcome of tariff reductions. Viewed in this light, the debate shifts to a more fruitful analysis of where scarce planning talents can be applied most usefully to integration. Beyond the establishment of a gradual process of tariff reduction and of institutions to compensate those who may suffer extreme hardships as a result of trade liberalization, for example, it is hard to see the need to plan in detail the integration of Latin American textile and clothing industries. Market forces can be expected to achieve integration in this field, with a minimum of inefficiency and public regulation. The same cannot be said

for the steel industry, if nothing else because of the presence of several public enterprises.

An ominous aspect of the insistence on planning integration in detail by the use of industry-by-industry complementation agreements and the accompanying reluctance to promote trade liberalization by across-the-board tariff cuts is a tendency to minimize the trade creation effects of integration, while promoting trade diversion. Granted that in the Latin American context some trade diversion may be an inevitable price which has to be paid for achieving noneconomic regional objectives, and that its economic cost can and should be minimized by planning, so as to avoid at the regional level the mistakes of national import substitution. But the benefits of trade creation in a regional or subregional common market could be substantial and are mistakenly dismissed by assertions of "lack of an industrial base" or of the unimportance of "static effects." A casual look at regional data, however, will indicate the relatively large share of industry in the Latin American GNP, as well as the large differences which exist among countries in the efficiency of industries which long ago totally replaced imports. Vigorous smuggling of Latin American products along several frontiers of the region also hints at possible powerful trade-creating effects of tariff reductions, even where infrastructure is deficient.

Regional planners could complement the impact of gradual trade liberalization, not only in concentrated activities, such as steel and petrochemicals, but also in the organization of regional research and educational facilities, as well as in the framing of joint policies toward foreign investors. Opposition to trade liberalization is often rooted in the fear that non-Latin American corporations will derive most of its benefits and that their investments will flow into a few relatively prosperous Latin American regions. Agreement on a common set of general rules regulating activities of those corporations in regional or subregional common markets and the creation of Latin American multinational firms would do much to speed trade liberalization. Further coordination of Latin American public enterprises in their dealings with foreign investors, especially in the field of natural resources, could also be useful. Planning in this area would serve not so much to improve the workings of the market mechanism but mainly to insure that certain noneconomic goals are not ignored in efforts to promote economic efficiency. The goal of geographically balanced development, together with that of greater Latin American autonomy, can probably also be best served at minimum cost to efficiency by the fram-

⁴ See, for example, Michael Bruno, "The Optimal Selection of Export-Promoting and Import-Substituting Projects," in *Planning the External Sector: Techniques, Problems and Policies*, Report on the First Inter-Regional Seminar on Development Planning (United Nations, 1967), pp. 88-135.

ing of common policies, with general incentives and penalties, rather than by blocking or regulating in painful detail trade liberalization.

Foreign Finance and National Planning

Presumably, while the effectiveness of aid is increased when given in the framework of plans, the continuity and stability of planning is enhanced by aid. Latin American experience during the 1960's provides only modest support for the latter proposition. The problem here is not just the weakness of the Latin American planning machinery, but also weaknesses and gaps in bilateral and multilateral aid.

Consider, from a national planning point of view, the predictability of concessional foreign finance coming in the form of project loans. The time elapsed between project preparation and its acceptance for financing by external sources can be not only lengthy but also unpredictable, for

reasons endogenous as well as exogenous to the recipient country. So is the lag between the approval of a loan and its disbursement. To illustrate the latter lag, and its variability, Table 2 shows average disbursement profiles of a large (112) sample of project loans made to Brazil, mostly during the 1960's. The loans are subdivided according to their sectoral destination. Besides the average rate of disbursement, expressed as cumulated percentages of the loan principal, its standard deviation is also given.

On the average, only half of the loans were disbursed about two and a half years after the signing of the loan agreement. Substantial variability in the disbursement rate may be noted among different categories, reflecting differences in gestation periods of projects (electricity versus others), as well as differences in institutional arrangements (many loans to industry were of the two-step variety, resembling program loans in the

TABLE 2
AVERAGE DISBURSEMENT PROFILES OF PROJECT LOANS TO BRAZIL
(As Cumulated Percentages of Loan Principals)

	All Loans	Agri- culture	Industry	Elec- tricity	Health and Water	Transport
<i>Disbursements by December 31 of:</i>						
Year of signature of loan agreement.	5.4 (17.1)	1.2 (3.1)	16.4 (30.1)	0.7 (2.3)	2.0 (4.7)	4.0 (8.2)
First full calendar year after agreement. . . .	21.0 (27.7)	20.7 (29.7)	35.0 (36.1)	13.9 (19.5)	19.5 (16.4)	14.5 (25.6)
Second full calendar year after agreement. . .	48.6 (34.5)	52.7 (33.1)	72.2 (34.5)	35.3 (27.5)	54.7 (24.9)	30.8 (34.3)
Third full calendar year after agreement. . .	72.9 (28.6)	76.4 (38.3)	89.5 (19.2)	65.1 (21.7)	81.2 (15.4)	38.9 (35.4)
Fourth full calendar year after agreement. . .	86.2 (22.0)	97.5 (2.5)	98.7 (3.3)	78.9 (20.4)	89.0 (10.2)	54.1 (35.8)
Fifth full calendar year after agreement. . .	95.6 (10.0)	97.5 (2.5)	100.0 (0)	93.7 (8.1)	95.1 (8.8)	100.0 (0)

SOURCES AND METHOD: Total sample includes 112 project loans; 9 in Agriculture, 29 in Industry and Mining, 37 in Electricity, 17 in Water, Sewerage and Health, 13 in Transportation, and 7 in others (not shown separately). Loans include those made by the Inter-American Development Bank (IADB), the International Bank for Reconstruction and Development (IBRD) and the U.S. Agency for International Development (AID). For the IADB, the calculation includes loans made from Ordinary Resources, Special Operations and Social Trust Fund from 1961 through 1968. Basic data obtained from the *Annual Reports* of that institution (Statements of Approved Loans). IBRD loans include those made during 1958 and 1959, plus those granted from 1965 through 1968. Basic IBRD data obtained from its *Monthly Statement of Loans*, December issues, mimeographed. Calculations for AID include only project loans made from 1962 through 1968; basic data obtained from AID, Office of the Controller, *Status of Loan Agreements*, December issues. Because many loans have been effective less than six years, the sample on which the averages and standard deviations (in parentheses) is computed becomes progressively smaller as one goes down a given column. *Ceteris paribus*, the slower the disbursement rate, the lower the grant element of a given loan.

speed of their first-step disbursements). High standard deviations also indicate great variability in disbursement rates within each category.

One lesson which planners (both in recipient and donor countries and institutions) may derive from these considerations is the importance, if project borrowing becomes an important source of external finance, of building up and maintaining a large pipeline of projects in different degrees of preparation and execution. Just taking into account the lag between the granting and disbursement of loans, and assuming that it takes on the average five years to disburse a loan after its approval, the steady-state pipeline will be about $2.5L$, where L is the desired annual inflow of external funds in the form of project loans. For most Latin American countries, the building up of such pipeline took place during the early 1960's. It is, therefore, a misguided criticism of aid to the region to point to a growing volume of authorized but undisbursed funds, or a growing pipeline, as evidence of low absorptive capacity. This, of course, does not mean that disbursement rates could not be speeded up in many cases by administrative improvements; the point is that even with those improvements disbursement lags will remain, requiring large pipelines as normal elements in the machinery for resource transfers. And when project aid starts from low levels, the pipeline build-up will be the first symptom of expanded aid intentions.⁵

The centralized preparation and/or selection of projects which can be externally financed is an expensive process and a large and diversified pipeline will be necessary to insure a fairly steady inflow of foreign resources coming in under project loans. Program loans, typically disbursed in

⁵ The Inter-American Development Bank, which started lending during 1961, presents an example of a pipeline build-up from zero, as shown in the following figures:

one or two years, have been used by AID in Latin America in part to correct those weaknesses of project lending. However, AID program loans have been given for one year at a time, and their main emphasis has been to support one-year fiscal and monetary policy packages rather than long-term development plans.

The fact remains that available aid sources do not provide Latin American planning with more than either short-run global support or medium-run support for a few subsectors. Looking at recent years, it is even questionable whether aid flows helped to make Latin American exchange receipts more stable. Simple measures indicate that Latin American export earnings during that period show greater stability by themselves than the sum of exports plus major gross aid inflows.⁶

following figures:

	Loans Approved Each Year (Million \$)	Annual Disburse- ments (Million \$)	Ratio of Accumulated Undisbursed Loans (Pipeline) to Annual Disburse- ments
1961.....	293.7	6.6	42.9
1962.....	329.4	58.7	9.4
1963.....	258.9	141.0	4.7
1964.....	299.3	198.1	3.8
1965.....	373.5	182.2	5.1
1966.....	396.1	211.8	5.2
1967.....	496.4	242.5	5.4
1968.....	430.8	290.9	4.8

Basic data obtained as in Table 2. Because of adjustments and cancellations, the accumulated difference between the first and second columns is not identical to pipeline figures given in the IADB annual reports.

⁶ Consider the following year-to-year percentage changes:

	All Latin America		Brazil	
	Merchandise Exports	Merchandise Exports plus Gross Official Disbursements	Merchandise Exports	Merchandise Exports plus Gross Official Disbursements
1962.....	6.5	0.8	-13.5	-22.1
1963.....	6.5	9.7	15.8	20.6
1964.....	7.4	4.5	1.7	-6.0
1965.....	5.0	5.4	11.5	18.1
1966.....	6.4	7.1	9.2	8.6
1967.....	0	-0.9	-5.0	-7.2
1968.....	4.6	6.4	—	—

There is no shortage of proposals to strengthen the contribution of foreign aid to national planning efforts. Many Latin American countries, however, could also take better advantage of external resources by improving their financial planning and their managing of the external debt. One possibility is to rely to a larger extent on borrowing from international bond markets, following the example of Mexico. While this source of finance is likely to be more expensive than AID soft-loans, its judicious use by the more advanced Latin American countries could contribute to the continuity and stability of long-term planning, avoiding the frictions which are often generated by the aid relationship. Similarly, greater attention by national planners to suppliers' credits could pay off, not only in the avoidance of periodic liquidity crisis and a better allocation of resources, but also in lower interest rates and softer conditions. Competition among industrialized countries eager to export goods and capital as well as the remarkable expansion of world capital markets can benefit the careful and organized shopper. Large semi-industrialized Latin American countries could do this shopping on their own; smaller ones may want to do it via regional or sub-regional financial institutions. Such regional or sub-regional groupings, organized to improve Latin American bargaining on trade and financial matters, will have a healthier footing than many existing inter-American institutions, where the U.S. presence generates a specious Pan-Americanism.⁷

Concluding Remarks

Both planning and aid are relatively new phenomena for Latin America. Difficulties and false starts could have been expected in the early ex-

perimental stages. If these early difficulties are to be overcome, planners must establish a proper balance between direct controls, general policies, and market forces in guiding resource allocation. The policy instruments are, in most cases, at hand. The need to use them, if nothing else to correct distortions introduced by misguided previous public policies, is clear in second-best Latin American economies.

The point can be generalized to other fields and, indeed, to other developing countries. The Bhagwati⁸ conclusions that "the fundamental thesis . . . is thus not that India does not need planning but that Indian planning is not as conscious of choices, alternatives, returns and costs as is necessary for efficient planning" and the desirability of "a form of planning where key, efficient decisions are taken in selected areas and the rest is left largely to the market" can be equally well applied to Latin America. Latin American planners would do well to note the Indian experience, where sophisticated planning tools have been in use for nearly two decades and where it cannot be said that planning lacked political support. There are so many areas where better planning is an obvious necessity in Latin America (education, improving the domestic capital market, public works, public enterprises, etc.) that it is a great waste to attempt detailed regulation of such matters as import demand, where market forces coupled with general public policies could save a good deal of planners' time as well as improve general economic efficiency. Note that these considerations apply regardless of the social system one considers ideal for Latin America; justice is not done to socialism by arguing that overvalued exchange rates and *ad hoc* protectionism are its harbingers. The resurgence of extreme forms of neoliberalism in Latin America has in fact capitalized on the errors of misguided statism and on the bankruptcy of its intellectual rationalizations. It is somewhat ironic, incidentally, that the Latin American country whose overall development strategy at the moment is most in accord with the classic principle of comparative advantage and which relies most heavily, in per capita terms, on net foreign finance, is socialist Cuba, a country which before 1959 saw its export market limited by the protective schemes of industrialized countries.

Skepticism was expressed earlier in this paper regarding administrative reorganizations as the way out of "the crisis in planning." Nevertheless, once a clear conception of the role of planning in

Data on gross official disbursements to all Latin America include all U.S. aid, plus that of the World Bank group, the IADB and the International Monetary Fund; obtained from Inter-American Committee for the Alliance for Progress, *El Financiamiento Externo para el Desarrollo de la América Latina*, Document OEA/Ser.H/X.14/Cies/1382 (Washington D.C., May, 1969), mimeographed, p. I-24. Brazilian data obtained from Document (of the same organization) OEA/Ser.H/XIV/CIAP/298 (Washington D.C., Oct., 1968), p. 198. Part of the instability in aid flows, of course, reflects political as well as economic instability within Latin American countries.

⁷ For several countries, foreign debt rescheduling and consolidation may be a necessary first step before greater access to world capital markets can be achieved. It may be noted that while the worldwide inflation of World War II and its aftermath cut the real value of Latin America's foreign exchange reserves (large at that time), that of the late 1960's has cut the real burden of Latin America's large foreign debt.

⁸ See the book review by Jagdish N. Bhagwati appearing in *Econ. J.*, Sept., 1969, pp. 636-37.

mixed societies is established, certain administrative changes could be found to be desirable. Public policies affecting foreign trade, for example, are taken in Latin America by several ministries and public institutions, often having little to do with one another. Even balance-of-payments statistics are sometimes estimated independently (yielding different results) by planning commissions, central banks, and the other *ad hoc* public bodies. Given the importance of the foreign exchange bottleneck, elementary good government requires a closer coordination among all of these institutions. In some cases, new EXIMBANK-type bodies could provide the platform for more integrated and aggressive foreign trade and

finance policies. Planning offices, in turn, could continue making balance-of-payments forecasts, but forecasts which are explicitly conditional on specific public policies. The quantification of the implications of different policies would then become the major contribution of these offices to rational policy making.

All of this assumes a strong public sector, relatively free from narrow special-interest pressures arising either nationally or from abroad, and interested in long-run development for the benefit of the whole society. This assumption is not valid for all Latin American countries, but several are close enough to satisfying it, so as to make previous remarks, one hopes, relevant.

FROM CAUSATION TO DECISION: PLANNING AS POLITICS

By STEPHEN S. COHEN
University of California, Berkeley

This paper analyzes the workings and potentials of French planning as a prototype model of modern capitalist planning. Its principal concern is methodological. How can we analyze, categorize, compare, and criticize planning processes?

The French planning process is not a streamlined design of smoothly fitting parts. Its formal structure tells little about its functional structure. Its explicit targets do not define its operational role. The plan is a collection of activities which have never been integrated into a single, coherent process. That is perhaps why there has been so much confusion about the way it operates; it operates in several ways at once.

The French plan has two principal components. Each is a complex system possessing a powerful logic of its own. Each is based on a different planning model and each model implies a radically different conception of the political function of planning. Each pulls the plan in a different direction.

The first component is a complex institution of daily, pragmatic state intervention in the activities of the major industries. The second is a formally coherent set of output targets—the general resource allocation plan.

I

The first component of the planning process (direct intervention) operates within what François Bloch-Lainé, France's foremost modern civil servant, calls the *économie concertée*: a close partnership between big business and the state to manage the rapid, but orderly, modernization and expansion of the industrial core of the economy. It is effective. The *économie concertée* (and that portion of the plan it subsumes) has been an important force in stimulating and steering the postwar development of the French economy. It works primarily through long-term actions on the supply side, where it seeks to change the directing attitudes and methods, as well as the structures and outputs of the principal industries.

The workings of the *économie concertée* are most visible in the modernization commissions of the plan, where businessmen, bankers, civil servants, planners, and "experts" assemble to prepare investment and restructuring programs for the industry. The initial appearance is one of quiet conflict. The civil servants have both objec-

tives and power. The objectives are concrete: investment programs, output targets, and, often, specific reforms such as a radical restructuring of the industry or a major export drive. Their powers match their objectives. They are discriminatory stimulants—selective subsidies to make a project desired by the state profitable to the firm. The businessmen expect to negotiate with the planners and civil servants; they do not expect to take orders. First, the giant firms and powerful trade associations that participate in the planning have their own direct influence at all levels of the state and these represent, in the rare case of a serious direct conflict, a powerful counterforce to the intervening civil servants and planners. Second, the fact that some 25 percent of the relevant group of civil servants will, before too long in their careers, switch into more lucrative employment with those same firms does a lot to obviate conflict. But business' most important bargaining card is the very nature of the planning activity. It is essentially a promotional program; negative control is weak. The planners and civil servants have the power to prod business to undertake projects. They do not have the same kind of discretionary power to impede a firm's expansion. Incentives are one-way affairs in dealings with large firms. Only with small firms can the withholding of subsidies act as a strong negative control. Furthermore, because big firms are the most actively involved, and big firms almost always have one or two projects on the drawing boards that fit well with the objectives of the plan, most of the conflict reduces to shadow boxing about the forms and extent of the incentives. These, in the process, lose much of their promotional character and become simply rewards for good behavior.

The conflict model misleads fundamentally. The essential *modus operandi* as well as the goal of the system is precisely the substitution of cooperation for conflict and competition. The *économie concertée* acts through and aims at cooperation instead of conflict between big business and the state—cooperation instead of competition among the firms in an industry and cooperation and coherence instead of confusion and uncertainty among interdependent industries. The French system of planning in which civil servants and the managers of an industry prepare detailed

programs for the expansion and modernization of that industry implies the further step of planning by the industry for the partition of that growth and development among its constituent firms. In brief, the component of the planning process circumscribed by the *économie concertée* is a system of cartels. But they are cartels with a difference. The goals are expansionist and modernizing, not restrictive and protectionist. And the state is an active, initiating partner, not a distant policeman. Its role is to create the structures of cooperation and through them to guide the economy towards expansion and modernization.

Cooperation is possible because it is based on substance—on a fundamental harmony of interest between big business and the state. Economic growth and modernization are the civil servants' highest priorities, and they are easily reconciled with the goals of the major corporations. The civil servants have no objections whatever to high profits, provided that they are part of a process of high investment for expanding production and increasing productivity. Developing the complex patterns of partnership took ingenuity, effort, and time, but it did not represent a radical departure from French tradition. There was never much enthusiasm for the practice or even the rhetoric of competition among French businessmen and only disdain for the noninterventionist role of the *laissez faire* state among the civil servants. The state had always "intervened" in the economy, and firms had always cooperated in ententes and cartels. The important and innovative development was the conversion of the protective cartel system into a modernizing, expansionist cartel system. A broad complex of historical forces—too broad and complex even to sketch in this brief note—was responsible for the development of the *économie concertée* mentality of cooperative modernization among businessmen and civil servants which, in the judgment of this writer, has been the single most important factor in the long-term success of the postwar French economy.

Participation in the *économie concertée* is as narrow as the goals. Because its efforts are directed at strengthening the industrial core of large firms, there is no need to obtain the active cooperation of a wide range of groups—trade unions, consumer groups, small business groups, peasant organizations and Parliament remain outside the system. They are not needed, either to choose objectives or to carry out programs. As the managers of the state and the big corporations see it, broad representation would bring in peasant and shopkeeper groups nostalgic for an

irrational past, trade unionists nostalgic for an irresponsible future, and politicians—all too eager to serve those groups. It would complicate matters, perhaps even destroy the system. As the range of concerns broadens to include a new sector, the management of its firms—whether they be private or public firms—can be brought in. And places can be kept at the conference tables for the "responsible" trade unionists the planned industrial evolution is supposed to produce. When they arrive, they can make positive contributions to the smooth management of the economy.

The *économie concertée* is based on a simple political ideology and defines a simple political role for planning. The state needs a high performance economy. This has come to mean a fundamental commitment by the state to the expansion and modernization of the big business sector. Big business needs the active cooperation of the state. It needs the state to maintain a high level of effective demand and to socialize many of its costs; e.g., the training of high-skilled manpower. It also needs the aid of the state in the management of its own affairs. The overarching organization provided by the state helps the industry to regulate competitive forces. The state also provides investment financing, technical assistance, and marketing information which individual firms cannot provide for themselves. In brief, big business finds that it needs the cooperative economy, and it needs the state to organize that cooperation. Most modern capitalist nations are developing some variant of the state-big business partnership model, but nowhere with such clarity and enthusiasm as France. The clarity is a noble French tradition; it needs no explanation. The enthusiasm is easy to understand.

From the perspective of big business, the *économie concertée* is the most satisfactory reconciliation of its potentially conflicting wants. It wants the active involvement of the state in the management of the economy, but it fears opening economic decision making to popular participation. Hence, the *économie concertée*, which makes planning into a device for providing big business with the active participation and support of the state while keeping broad participation politics away.

Under the slogan of keeping politics out of planning, the *économie concertée* component of the planning process confines the plan to the special politics of economic administration by the managers of big business and an elite corps of civil servants. Within that role, the plan is effective. But the range of its effectiveness is limited to those areas of economic decision making

that can be incorporated into that special political subsystem—essentially to the long-term modernization of the supply side. Important areas of economic policy and activity, however, lie outside the boundaries of the role for planning established by the politics of the *économie concertée*. Short-term aggregate demand policy is one. The exercise of countercyclical policy is the focal point of a very different politics of broad participation and conflicting interests. The incongruence of the two political processes is the principal reason for the complete failure to coordinate short-term demand policy with the longer-term supply side programs.¹

Until recently, the plan has preferred to sacrifice major influence on short-term demand policy—and avoid the profound political changes in the planning process that such an influence would necessitate. It remained within the rather comfortable confines of the *économie concertée*, seeing its essential task as the promotion of long-term, supply side modernization. But a new set of forces is pushing the plan beyond the boundaries of the *économie concertée*. The most important force is simply the dynamic of its own success in helping to create the structures and attitudes for self-sustaining modernization within the big business core of the economy. The central economic problem is no longer one of developing and modernizing supply side productive capacity. Instead, it is becoming one of how to handle the growing interdependence and complexity of the newly modernized economy. As the institutional structures of the economy change and become more integrated decisions affect more and more people, institutions and groups, and more and more groups, institutions and people seek to affect each decision. As conflicts broaden and can no longer be contained within traditional institutions, new, more centralized institutions are forced into being. This is clearly seen in conflicts over income distribution which are increasingly settled through centralized political bargaining instead of through decentralized markets. The government is increasingly called upon to balance distributive claims of different groups, and finds itself developing an incomes policy. The need for incomes policy is further intensified by the opening of the French economy into the Common Market. This combination of pressures is pushing the plan into a direct concern with price planning and incomes policy. To accommodate these new kinds of prob-

lems, the plan will have to change fundamentally. The plan has been functioning in a political role, according to a logic defined by the political system we have called the *économie concertée*. Its activities can only be understood in terms of the politics of the *économie concertée*. But the politics of the *économie concertée* cannot accommodate the new range of concerns and the new kinds of problems that the plan is being forced to confront. The plan is being forced beyond the limits of the *économie concertée* toward a new political role. That new role is implicit in the second but noneffective component of the plan: the general resource allocation plan. The substantive political content of that role is the principal reason the general resource allocation component has remained ineffective.

II

The second component of the planning process is the general resource allocation plan. Its technical basis is familiar to economists: an input-output table; a set of input coefficients; a set of estimates for final demand; and a set of primary supply constraints. The problems associated with this approach to planning—accuracy of input coefficients under conditions of changing prices and technologies; appropriateness of levels of aggregation; reliability of demand estimates, etc.—are also well known. Its political logic, however, and its implicit political role have been less thoroughly discussed. That is our purpose here. For implicit in the basic characteristics of general resource allocation planning is a role for planning that contains a potential for important innovation in the political process.

The general resource allocation plan pushes planning towards the center of broad participation, competing-interest politics, towards precisely those forces that the *économie concertée* avoids. The thrust is generated by the basic characteristics of the formal information system that constitutes the "technical" basis of a general resource allocation plan: consistency and comprehensiveness.

Consistency is primary. Its technical meaning is simple. When each target is a function of all other targets, each target makes sense only in terms of the planned pattern. It follows that the rationality of the particular quantitative value attached to each target rests on the assumption that the entire pattern of planned targets will be realized without major distortion. The political meaning of consistency is quite as simple: in both its preparation and its implementation, the plan is a package deal.

¹ Many of the arguments and themes sketched in this paper are properly developed in my *Modern Capitalist Planning: The French Model* (Harvard Univ. Press, 1969).

It is a very big package. The planned pattern must be comprehensive as well as consistent. All economic activities must be accounted for so that the partial equilibria represented by each of the targets add up to the aggregate balances for manpower, foreign exchange, investment and savings that define general equilibrium. This does not, of course, mean that the plan must contain explicit, numerical targets for each and every identifiable economic activity. All it must do is account, directly or indirectly, for all activities. Fixing numerical output targets for such products as sweaters, sunglasses, or tennis shoes is neither necessary, desirable, nor possible. Instead, broad product groupings can be established and targets assigned to cover the aggregate impact of those activities. Just how resources are allocated within the category is not important—so long as the totals hold. Similarly, certain questions which stand as the logical center of the plan can often be treated indirectly. For example, the French plans have made no direct statements about income distribution. But detailed assumptions about income distribution are necessarily incorporated into the projected pattern of final demands from which the rest of the plan is derived. In brief, comprehensiveness does not mean exhaustive enumeration; it means all-inclusiveness at some level of aggregation and some degree of directness.

Because the plan must incorporate, directly or indirectly, targets for all major economic activities, it becomes a framework for political decisions. It cannot simply ignore the impact of military expenditures, or farm policies, or foreign aid, or automobile production, or highway construction, or tax incidence. But broad participation, conflicting interest politics focuses on precisely these decisions. And once the planned pattern, incorporating the major decisions on resource allocation is established, it must be implemented *in toto*. Technically, distortion would destroy the rationality of the output targets. Politically, distortion would mean that the bargaining that went into the preparation of the pattern of targets was a sham.

In order for the planned pattern to be realized, all the major collective decisions embodied in its targets must be carried through. The word "decisions" is used to distinguish the independent from the dependent variables, the initiating from the responding forces. Only if the key "decisions" are carried through as planned will "market forces" steer the rest of the economy to the planned targets. The most important of these decisions concern the public sector.

The government must commit itself in a de-

tailed fashion for a long time. During the planned period it cannot launch major unplanned military projects, or space probes, or welfare programs, or tax reforms without destroying the plan's coherence.² It must coordinate its use of short-term economic controls with the middle-term objectives of the plan. The commitment to planning cannot come from the government alone but must come from the major social forces. Either the various political elements—big business, the trade unions, small business, peasants, the military, the permanent bureaucracy, etc.—will believe that they can gain satisfaction through the plan and begin to focus their efforts on the choice and execution of its objectives, or they will refuse the politics of planning and continue to concentrate their energies on influencing the exercise of short-term policies and on initiating pragmatic programs. The latter course would result in continuing the traditional pragmatic-compromise method of policy formulation and execution; i.e., politics as usual. Whatever decisions had been previously written into the plan would not be carried out—at least not without the kind of modification that would distort the plan. The third alternative—plans formulated and efficiently implemented without political participation—is, of course, tyranny. Given the necessarily comprehensive nature of the plan, totalitarianism is perhaps a better word. In order to be rational and effective, the plan has to be comprehensive and consistent in both its preparation and implementation. For that to happen, the plan must become a principal framework for political discussion and the planning process must become a principal political arena. In brief, in a democratic society, a general resource allocation plan can be rational and effective only if it is the output of a broadly participatory political process.

The basic characteristics of a general resource allocation plan push it into the center of broad participation, conflicting-interest politics. They also indicate the likely effects planning would have on politics.

III

The politics of a general resource allocation plan are comprehensive, simultaneous and explicit. The plan centralizes both political space and time. It pulls a vast range of decisions into a single framework and insists that they be decided simultaneously. The politics of frequent, individual, and partial decisions become one of simul-

² A "revolving" plan model would help to relieve some of the pressure, but it would not change the fundamental political logic.

taneous, interrelated, and long-term decisions. The traditional interplay of issues, interests, and marginal adjustments must accommodate a new politics of social design. For the various political groups this means abandoning the traditional politics of frequent, piecemeal gains and losses for that of single, long-term package deals.

Of the three characteristics of the politics of general resource allocation planning—comprehensiveness, simultaneity, and explicitness—explicitness is the most important. The fundamental output of a plan is explicitness. The plan makes explicit the implications of choices among alternatives: the greater the range of implications anticipated (one can just as easily say planned), the more sophisticated the planning. Explicitness is what is essential to planning. Efficiency is not; it is a use to which the explicitness of planning can be put. And it is not the only use.

The explicitness of a general resource allocation plan constitutes a new information system. It provides totally new kinds of economic information, which are necessarily distributed differently from the way economic information is distributed without a plan. It provides—to the public at large—a comprehensive and rather detailed picture of the direction of economic development.

A clear picture of the pattern of future development is simply not available under the present system. By providing such a picture the plan makes possible political discussion about the shape of development. Such a shift in political discussion—from disjointed, partial questions and projects to alternative social designs—would constitute a major innovation (and elevation!) in the political process.

By increasing the amount of knowledge in the system, the plan increases potential power in the system. The range of potential environmental control—and therefore choice—expands. Shaping the future of society passes from a process of causation to a process of decision. At present, the direction of development is “caused”; it is not decided. In a market system the shape of development is caused by complex interactions of countless decisions that are all taken without any concern as to the shape of the whole: the pattern of development is an unintended outcome. Pragmatic politics enters the causation process in disjointed, piecemeal fashion to temper and to steer it a bit. But it does not substitute an alternative process of deliberate decision for the present process of unplanned causation. General resource allocation planning implies such a shift in the allocation process.

But the shift in process implies a shift in sub-

stance. Different methods of formulating and expressing objectives generally yield substantively different results. As the deliberate decision process of planning overcomes the causation process as the principal mechanism for translating social wants into effective demands, the range of choice changes. Important social objectives such as the preservation or transformation of particular social structures, which had difficulty finding direct expression through the market mechanism, may become more immediate and more effective forces in allocating resources. The substantive content of development will change. There will be shifts in the distribution of rewards and major changes in the distribution of economic power.

Because the change in process (from causation to decision) implies such major changes in substance (in incomes, profits, prestige, freedom, power, etc.), it is not likely to happen without a prior, enabling change in substance. A major shift in political power is necessary to change the allocative process.

IV

Romantic reactions—whether they call for free market and invisible hands, self-sufficient rural communes, or *Gemütlichkeit* old neighborhoods—lead to social disasters. France's *économie concertée* is not the only nonexplicit, nondemocratic effort at planning. It is merely the most visible—and one of the most competent. An informal partnership between big business and the state dominates the central industrial core of most modern capitalist economies—even this one. Major efforts at planning are to be found at all levels—from the macro “fine tuning” that is causing so much static at the present moment to micro planning by giant corporations. These have developed in response to the changing institutional structures of the economy, the growing recognition of interdependence, and, most important, the growing recognition of the ever more exorbitant costs of nonplanning. Furthermore, the “problems” that now confront the economy—the acute manpower problem of a black subproletariat and a rapidly changing job mix, the environmental problem of profuse pollution, the problems of conglomerate corporations and agglomerate cities, etc.—will all accelerate the push toward planning, toward the deliberate direction of economic forces. Realistically, the economy faces not a question of planning or not planning, but rather a question of what kind of planning.

Apolitical planning is a fiction. All planning is political. The nature of a particular planning process is inseparable from the substantive outcome of

the conflicts it "treats." It can never be neutral. Each planning process defines, explicitly or implicitly, a political role for planning; it makes planning part of a particular kind of politics. The "One Best Way" approach that dominated city planning behind the slogan, "Keep Planning Out of Politics," was political planning. It kept planning out of one kind of politics and confined it to another. This paper has tried to show that the French *économie concertée* is just as political as the explicitly politicized general resource allocation model, but that the *économie concertée* planning model

belongs to a very different kind of politics and can only be understood as part of that special politics of narrow-participation economic administration.

Does conventional economics provide a mode of analysis suitable for analyzing, categorizing, comparing, and criticizing planning processes? Can it relate planning process to political substance, political power to economic techniques? In brief, can it speak to the question, What kind of planning? In the recent past it has not, and the momentum of its present concepts, demarcations and models gives evidence that in the near future it will not.

PLANNING WITHIN THE FIRM

By JOSEPH L. BOWER
Harvard University

It is said of some business schools that they are prepared to disqualify a man as potential faculty material if he happens to be proficient in economics. It may be said of most economics faculties that they are prepared to disqualify a man as potential faculty material if he happens to know very much about the firm. It is a hallowed tradition dating back to the religious order in the fourteenth century that tried to determine the number of teeth in the head of a horse by interpreting the Bible. They disqualified the novice who brought in a horse for study. While sanctioned by time, attitudes such as those described above make scientific progress difficult in the development and testing of hypotheses about planning in the large corporation. The problem is nowhere more evident than in the Galbraith versus Solow debate where two guys told it like they thought it was: "If I was a firm, how would I plan?"¹

The problem is that a large firm is one of the most sophisticated institutions for the conduct of organized purposive behavior ever invented. The appealing analogies are too simple and too intellectual. A large corporation is not a mutual fund. And plans cannot be evaluated like journal articles. In a large corporation, people must be moved before dollars can be shifted, and business plans and investment projects must be generated before capital can be invested in them.

I want to argue in this paper that "planning" is, for practical purposes, a complex process which, in addition to intellectual activities of perception and analysis, involves the social process of implementing formulated policies by means of organizational structure, systems of measurement and allocation, and systems for reward and punishment, and, finally, involves a dynamic process of revising policy as shifts in organizational resources and the environment change the context of the original planning problem. The outcome of what I shall call "the planning process" is the choice of major markets and products to serve them, and the commitment to allocate resources to the implementation of planning choices.

For purposes of making the argument, it is

convenient to use the paradigms of Galbraith and his reviewer, Solow, as foils. In an amusing but relevant discussion of the *New Industrial State* in the fall, 1967, *Public Interest*, each is particularly eloquent in presenting a view of the planning process in corporations that is inadequate for the purpose of formulating policy. These errors would be unimportant were it not that misunderstanding the strengths and weaknesses of the contemporary American corporation leads economists to overestimate the responsiveness of firms to planning in the design of policy proposals.

Galbraith. Let us look first at Galbraith's picture of the American corporation, and then turn to Solow's, in part because their dialogue developed in that fashion. Galbraith suggested that:

As viewed by the industrial firm, planning consists in foreseeing the actions required between the initiation of production and its completion and preparing for the accomplishment of these actions. And it consists also of foreseeing, and having a design for meeting, any unscheduled development, favorable or otherwise, that may occur along the way.

An obvious and critical question that one would ask is, "Planning for what? What is the firm trying to accomplish?"

Galbraith argued that the firm and its managers, by a process of mutual adaptation and identification, develop as a goal the pursuit of economic and technological growth. The technostucture pursues its own well-being as a goal:

The first requisite for survival by the technostucture is that it preserve the autonomy on which its decision-making powers rest. . . . Power passes to the technostucture when technology and planning require specialized knowledge and group decisions.

The marketplace does not police the technostucture's choices, says Galbraith, because firms use their power to shape the market rather than the reverse. By adaptation, the autonomy of the technostucture becomes a social goal.

Solow. Robert Solow has responded to Galbraith's argument by suggesting that firms produce, price and invest "as if" they were profit maximizers. More precisely, Solow argues that after one acknowledges the size of large corporations, the extent to which control is separated from ownership, and the fact that management is spread over a bureaucratic hierarchy, it remains

¹ Robert Solow, "The New Industrial State"; J. K. Galbraith, "A Review of a Review"; and Robert M. Solow, "A Rejoinder"; all in *The Public Interest*, Fall, 1967.

true that:

For any given amount of invested capital, a corporation will seek the largest possible profits in some appropriate long-run sense, and with due allowance for cheerfulness. If the return on capital thus achieved exceeds the minimum required yield or target rate of return, the corporation will expand by adding to its capital, whether from internal or external sources.

The issue, Solow argues, is "whether the art of salesmanship has succeeded in freeing the large corporation from the need to meet a market test, giving it a decisive influence over the revenue it receives." And later, "the real question is whether there is some other goal that business pursues systematically at the expense of profits." The evidence, Solow argues, is that the market is still a primary constraint on business behavior, and that no other goal is pursued "systematically at the expense of profit."

My argument is that neither Solow nor Galbraith comes very close to describing the behavior of real firms. That is why each has so much unexplained behavior to account for. In fact, the research which they cite is largely aggregate studies of investment data, or the results of various other large sample survey research. Galbraith's reliance on the writings of Herbert Simon, James March, and Richard Cyert does not provide much help either, because none of their work is based on systematic study of the general management of large organizations.

Solow wins the debate on points because "there is," as he notes, "a certain amount of positive evidence that supports the hypothesis of rough profit maximization." But that evidence is really at the level of what he describes as a "detectable tendency." In fact, the measures of economic activity that are used as evidence are extraordinarily crude, and Solow joins Galbraith in attributing gaps between prediction and observed behavior to stupidity—except that Solow prefers the word "obtuseness."

I must apologize to my audience here because I am about to drag in the horse. There really is a good deal known about the way in which large firms plan, produce, invest, and set prices. A number of researchers have studied the planning process in firms first hand and reported their findings. Rather than break up the text with footnotes, I want to acknowledge that the description which follows is based on the work of many.

A Look Inside the Large Corporation. Let us examine what we already know and how it differs from Galbraith's and Solow's position. The most critical aspect of the large corporation is the de-

gree to which it is specialized, and the extent to which those specializations are reflected in subunits of the organization. In order to accomplish the many tasks involved in conducting business efficiently, the large corporation has organized its physical and human resources to provide dedicated facilities for each of the many kinds of activities in which it is engaged.

Often the large corporation is organized to provide specialists for planning and implementing all of the following activities: the capture or purchase of raw materials at major raw-material sites; the management of production facility locations (which may have been designed for the production of many products); the process for producing products (the same product may be produced by more than one process and at different facilities); the development and management of a product line; the analysis of opportunity in the industry markets towards which products are directed (many products may be produced for the same industry); the sale and distribution of products to customers (who may buy in several industry markets); the sale and distribution of products by customer location (national customers have offices in many parts of the country and it may be economic to sell to them on a regional basis); the service functions, such as accounting and law and especially engineering and research; the coordination and management of increasingly aggregate groupings of these specialists.

It is worth noting that there are now special kinds of general managers. Some coordinate specialists. Others evaluate first-level generalists, and integrate the long-range plans of subunits. Still others focus on selecting the middle-level generalists and on modifying the structure within which the firm plans and operates.

In order to provide for the systematic application of specialized capability, a typical corporation is organized into a large number of subunits. There are sections of plants, research departments, sales offices, product managers, and an endless number of other kinds of units. Typically the corporations will have as many as seven levels of management above the operating levels; for example, section, department, division, group, corporate staff, chief executives, and the board. Specialists may report to generalists in the same area of specialization at several levels. Thus there may be a marketing staff reporting to a department general manager, a different marketing staff reporting to a division general manager and still another marketing staff reporting to the president of the corporation. The explanation in each in-

stance is an attempt to exploit specialized skills economically.

It is very important to realize that these organizational complexities are not merely administrative curiosities with passive consequences. They are not constraints, but instead, are active elements of the planning problem. The point is that a basic prerequisite for any kind of systematic rational planning is various kinds of quantitative information provided as an input, and quantitative measures of results provided as an output. Each of these subunits requires specialized information and specialized measures. If, for example, the subunit in question is a profit center of the division, responsible for production, engineering, sales, and research for a particular group of items, it needs data on demand, and costs, as well as trends to be expected. Typically, data on sales comes from another subunit—often a pooled sales force. Sometimes more than one sales force is involved—the usual case, for example, when there are exports. Cost data come from plants which may or may not be under direct supervision of the profit center. And data on competitors' costs or technology are likely to come from the sales organization, or engineering, or research.

Moreover, data describing inputs to one subunit often describe outputs of another subunit of the same organization. One man's profit is another's cost. Transfer prices help to reduce the interdependency of subunits but they introduce an element of bargaining into the information system. Even worse, every time the strategy of the organization shifts, the transfer price has to change. (What was a by-product may now be a primary product, or a free good.) Some managers may recognize the problem, but accounting is almost always a centralized corporate function, and getting a transfer price changed takes time. In the interim, a profit center manager must explain to his superior why the data he is using are inappropriate.

But far more aggravating to the problem of acquiring relevant information is the problem of time span, the problem of intelligence, and the problem of control.

Time Span. It is now well recognized in the discussion of corporate objectives that large firms seldom attempt to maximize that index of performance known as annual profit. At the very least, they are concerned with the economic prosperity of the firm over three to five years. In fact, students of business, including some economists, have recognized that true economic success has to be measured over long-run periods. The annual accounting profit looks, more or less, like the ac-

countants want it to look. In contrast, the results of a strategic move, particularly a major investment, are not known for years after the so-called "decision" has been made. And accountants can do very little to distort the extent to which outcome has varied from expectation.

It is well known in business that a good hatchet man can cut 10 percent to 20 percent of a given year's production costs out of almost any business. What is rare is the man who can cut costs significantly without damaging the ability of the business to function in succeeding years. "Costs," after all, are often "investments" improperly labeled. But the cost—that critical input to planning—looks the same on paper in either case. And generally the time span of any of the sources of information available to a manager is usually much shorter than the time span he would like to take into account when planning major economic commitments.

Not only are reported quantitative data imperfect measures of economic performance, but precisely for the same reasons they are imperfect measures of a manager's performance. If the results of a manager's actions cannot be known for years, the annual number is not a true measure of how well a man has performed. On the other hand, managers are paid, salaries are raised, and promotions made on an annual basis—or at best, on the record of the annual measures.

Intelligence. A second problem with data is the difficulty firms face in obtaining information describing their competitors. If a manager is planning to build a new plant to manufacture a product such as ammonia or vinyl chloride, he needs to know what capacity competitors will have and what their costs will be. The kind of error in a cost estimate which is trivial in a statistical cost study—say 10 percent—can threaten a producer of fertilizer or industrial chemicals with disaster. One cent per pound on a ten cent per pound item may be half the gross margin and all of the profit. It has been common in recent years for entrants into the fertilizer field to see their plants made economically obsolete before they came on stream. I am not clear as to how Galbraith would regard these competitive failures. Certainly what are catastrophes internally are seldom perceived externally as failure.³ Perhaps the results of business decisions viewed from the outside look more convincing because offsetting mistakes, delays, corrections, and crash programs even out some of

³ And those catastrophes that do take place are more likely to appear in *Business Week* than in banner headlines. A new process to vinyl chloride isn't news, even if it permits Dow to push Union Carbide out of the market.

the unbalanced effects of imperfect decisions made under great uncertainty that radically affect the careers of individual managers.

Within each firm managers do their best to protect their personal stakes as they commit their firms to particular positions. By hedging personal bets and avoiding uncertainty where possible, they narrow the range of the problem facing them to a set of issues that can be resolved with the intelligence available. Even so, on a regular basis, shifts in market position do take place which—while they do not always change the ranking in the Fortune 500—often change the careers of the men responsible for the specific market in question.

It is acceptable to describe the competition among giants as rivalry rather than competition. From certain points of view, size makes firms seem much less responsive to the market than the word competitive implies. But it is wrong to ignore—as Galbraith does—the internal effects of rivalry on planning and on resource allocation. Plans are responsive to the stakes of planners, not merely the stake of the firm in the market.

Control. The problems of specialization, of coordinating multiple subunits, of the time span of measures, and of gathering and using intelligence all are reflected in the problem of control. The same data used for planning are also used for determining how well an organization is functioning. Every forecast may have the quality of an estimate, a target, or a commitment. A manager may want to forecast a cautious view of the future, but choose a posture of unadulterated optimism instead because he knows that otherwise he will lose.

In fact, the most ironic aspect of the measurement and information problem is that feedback from the organization is much faster and unambiguous than feedback from the market. The systems used to motivate and develop managers have short time spans. The systems used to control performance have short time spans. The result is that the data contributed by managers tend to be biased by short-run considerations.

In summary, the way in which top management organizes to shape the stakes of the planners critically influences the plans it will have a chance to approve. Consider the following example reported by the assistant comptroller of a divisionalized company with highly centralized financial control:

'Our top management likes to make all the major decisions. They think they do, but I've just seen one case where a division beat them.

I, received for editing a request from the division for a large chimney. I couldn't see what anyone could

do with just a chimney, so I flew out for a visit. They've built and equipped a whole plant on plant expense orders. The chimney is the only indivisible item that exceeded the \$50,000 limit we put on the expense orders.

Apparently they learned informally that a new plant wouldn't be favorably received, so they built the damn thing. I don't know exactly what I'm going to say.'

Now the most important point to be learned from this story is that things may not always happen as "planned." But once that elemental point is grasped it is also apparent that out of context, we have no way of really knowing why the managers in question violated the spirit of the corporate regulations regarding capital budgeting. (In some parts of the world, we could conclude that relatives of the key managers in question were in the construction business.) The proper interpretation depends on the motivation of the managers in question. The motivation of managers is critical to the problems of planning and control, for it determines in large measure how managers perceive and respond to opportunity and uncertainty.

Galbraith's notions of motivation in the technostucture are not adequate. As I will try to argue later, they are descriptive of behavior in the least interesting parts of the business sector. Consider by way of contrast the observations of Zaleznik, based on his research with managers:

Business organizations also serve as the stage upon which the conflicts of individuality are played out for many people. On the one hand, there seems to be a demand for conformity and identification with organizations that threatens the very essence of individuality—the sense of one's personal impact on events. On the other hand, organizations provide ample room for individuals to assert themselves and express their unique style of performance. The point is that organizations do not provide individuality as a gift. It has to be gained and even fought for while sustaining one's involvement and responsibility.

The crux of leadership is the acceptance of responsibility—the idea of fantasy that one can make a difference in the course of events. This sense of personal involvement in life is not simply a passive experience. It is an impelling urge to make a difference and use oneself in effecting outcomes. The insecurity of leaders is often related to the possibility that their actions in the end may appear trivial.³

In other words, when managers in large corporations move toward action they face personal and organizational challenges in substantive guise.

³ Quoted in Bower.

⁴ A. Zaleznik, *Human Dilemmas of Leadership*, p. 3.

The substance plays a part, but not necessarily a central part, in the outcomes of negotiations among managers that represent planning and resource allocation choices.

The implication of these findings is that when a large highly specialized bureaucratic hierarchy tries to cooperate in adopting an appropriate posture for the future, it must do so in the context of an organization chart, a system of measurement and information, and a system of managerial reward and punishment that primarily reflects the short-run needs of day-to-day operations. This context produces forces that increase the dimensionality of plans and the inputs to plans in ways that can be quite harmful for a company. Plans are not mere analytic manipulation of economic data. A plan is an argument for something a network of managers wants to do because it will be in their interest to do so. The job of top management is to keep self-interest and corporate interest aligned. The result is that good top managements spend most of their time modifying the context so as to improve it, calibrating the judgments of the men who man it, and selecting the best men to fill key slots in the future. These processes are the heart of top management in large firms. They are the processes both Galbraith and Solow ignore. And they are critical to how planning operates in large companies. Against this background, let us see what research has revealed about planning.

The Planning Process. In discussing the planning process, the first point to note is that the process is not focused at the top or the bottom of the hierarchy. The specific content of plans emerges from lower level subunits concerned with specific markets defined in terms of product and consumer. But the overall relationship of the company to its environment, including factors as the availability and cost of capital, negotiations with labor and most units of government, are typically the concern of top management. Middle levels of management face the difficult task of reconciling the multiple product-market plans of subunits with each other and with top management plans for the place of the company in the global environment.

A second aspect of planning is that the process shaping the content of plans—both the choice of objectives and the discrete commitment of resources—is different from the process that leads the plans to be approved. The former process is something like textbook idealizations of planning. It is concerned with technical and economic substance of a business or an investment. The process of developing the content of plans departs from idealization, however, for two reasons.

First, a subunit manager plans within the scope of his job as it has been defined and as he is measured. He is usually directly responsible for only part of the complex of specialties that constitutes a given corporation's participation in a product-market area. For example, a new facility is planned by the subunit responsible for productive capacity only when information from another subunit indicates that a discrepancy exists, that "costs are too high," or "sales (or forecasts of sales) exceed capacity." If a manager has too much to do, if the scope of his job is too great, he allocates his time in response to his crises. These may not be the same as the corporation's. If he has a lot to do, he may struggle to make this year's budget because he is measured on it and waste \$10 million on a poorly studied capital project because he is not measured on it.

A most important implication of this finding is the importance it places on feedback from the market (Solow not Galbraith!). One of the most highly visible phenomenon in a large corporation is when a product loses position in the market—sales or market share—to competition (rivals if you prefer). A subunit manager will always respond so as to try to avoid so apparent a sign of failure. This is one instance in which, with virtually any system of measure and reward, management self-interest and corporate interest tend to coincide.

A second point to note is that unless the scope of his job is changed or the measures, a manager will not usually change his behavior. If corporate policy of any sort—labor, location, tax problems, race relations, or finance—is to influence a manager's plans, explicit steps have to be taken by corporate management to modify the measurement of the subunit manager's job to take account of the policy. (Clean air is a great corporate objective, but meaningless unless plant managers are forgiven the higher costs involved.)

Let us turn now to the process that leads plans to be approved. It has been found to bear no resemblance to traditional textbook description. The treatment of a plan and the rate of progress of a capital project up the hierarchy of management depend entirely upon the way in which the different managers at the several levels of the organization interpret what the corporation wants of them, and how the plan or project is likely to reflect upon their performance.

The comments of a middle-level manager of a large corporation responsible for a group of products with sales of \$75 million illuminates the way in which this process works:

What it really comes down to is your batting average. Obviously anything cooked up, I have to sell and

approve. My contribution is more in the area of deciding how much confidence we have in things. The whole thing—the size, the sales estimate, the return, is based on judgment. I can kill or expand a project based on my judgment. I decide the degree of optimism incorporated into the estimates. You know your numbers change depending on how you feel. The key question is “how much confidence has the management built up over the years in my judgment?” A guy in my position must think this way. He loses his usefulness when he loses the confidence of higher executives in the company. Otherwise his ideas won’t be accepted when he goes up.

I can lose that confidence by being too optimistic or too pessimistic. I can lose the confidence of the people above, and I can lose the confidence of the people below that they, and the degree of confidence elicited from them, are not going to be ill used.⁵

Put more formally, by the way in which the careers of general managers are advanced or retarded, top management makes very clear its attitude toward the quality of judgment exercised by division general management. In turn a general manager sponsors a project when he believes it demonstrates the right sort of judgment and, therefore, will be in his interest to do so.

Once a project is sponsored, it is almost always approved by top management. They are loathe to second-guess the judgments of the men selected for intermediate-level management precisely on the basis of their ability to evaluate the technical-economic content of product-market subunit plans and projects. That is why batting average is so important. It reflects the ability of middle-level managers to judge lower-level generalists. It is also why trade-offs in the classical sense are so hard to make in corporate terms (explaining, in fact, some “nonmaximizing” behavior). Middle-level managers do not have a corporate perspective, and the corporation has delegated the power of choice.

Notice that this is not exactly what Galbraith describes as the vestment of power in the technostructure. Substantive competence rests at lower levels, but through control over the form of the organization and the elements of the reward system top management does influence the choices of technostructure. The social-psychologist Raymond Bauer has termed this process “meta-management.”

Notice, also, that the problems with data and measure discussed above will have a critical influence on the way in which performance is defined and evaluated. While everyone realizes that long-run return is the object, the only credible quantitative data is short run, so that ob-

served behavior usually consists of various sorts of short-run suboptimizations and/or a limited number of major moves justified by judgments of long-run strategic consequences. Both kinds of decisions reflect a detectable tendency to profit maximization, if viewed from sufficient distance—Solow’s point. But neither result from the traditional calculus of economic maximization.

By emphasizing one or two points from the preceding discussion, it is possible to bring this stage of a discussion to a close. Note (1) the content of plans is shaped by the structural elements of the corporation as they apply to product-market management; (2) the approval of plans is shaped by the same structural elements as they apply to intermediate levels of general management; (3) top management of American corporations—unlike the President and unlike top management of foreign firms—can and does manage the rules of the game in order to manage outcomes by an indirect process.

A New Order of Complexity. Before Solow gets a chance to say, “I told you so,” let us move on and look at the horse more closely. Up to this point, the discussion has proceeded as if all large companies were alike. In fact, they vary substantially along lines of geographic diversity (single location to multinational), technological complexity (computers to soap), and product-market variety. The differences are considerable in all three areas, but in this paper it should be sufficient to argue the importance for policy of only one extreme contrast.

Consider on the one hand a large vertically integrated manufacturer of a basic commodity such as paper. Consider, in contrast, a managed conglomerate such as General Electric. It should be distinguished from the financial conglomerates which are little more than holding companies. Certain key differences in the planning process should be apparent on inspection.

1. For the integrated paper firm selling a single group of products to a single major market, product-market is company-environment. Hence top management is involved in all phases of the planning process. For strategic purposes the firm acts as a single profit center, no matter how large it may be. Top management initiates the planning for new paper machines (\$40-\$50 million) or for acquisition of new forest acreage and they worry about the balance between raw material, intermediate, and final conversion divisions.

2. Because top management of such firms is involved in the initiation of strategic plans, they become deeply committed to them. Perhaps because it is hard to measure results of personal decisions critically, the managements of such firms

⁵ Bower.

have often been found reluctant to evaluate major investments or resources commitments against opportunities offered by other industries, and sometimes unwilling to consider departures from traditional approaches to doing business in their own industry.

3. There is a noticeable tendency for individuals to make their entire career in a single basic industry. Such men value their product in absolute rather than economic terms. They like forests and paper. Output is valued for its own sake, not for return on investment. Sometimes the only measures of performance available are physical; e.g., tons/day.

4. Because top management's role in big decisions is so important, and because of the great interdependency among functions and divisions it is hard to measure the contribution of subunit managers. Rewards tend to be based on technical proficiency and seniority rather than economic performance. These attributes almost define social standing in the corporate community.

In short, these vertically integrated firms really look a lot like Galbraith technostuctured corporations. Such firms often pursue sales and growth in the manner Galbraith attributes to technostucture, but here again Galbraith does not have it quite right because it is the involvement of top management rather than its passivity that seems to account for the observed behavior.

Galbraith discusses diversification as an aspect of the industrial state. But product-market diversified companies look and behave very differently.

5. In managed conglomerates, even those not as highly diversified as General Electric, plans and capital projects are initiated at subunit levels. Top management focuses on broad questions they perceive to be critical. A problem situation such as those posed by the computer business or the SST constitutes the only occasion on which the top managers will become deeply involved in the substance of strategic questions.

6. In the diversified corporation, the divisions are measured more by profit, or return on investment. Managerial rewards are more closely tied to quantitative measures of economic performance as opposed to social standing in the society of the corporation. In effect, the close relationship between social and economic systems is stripped away leaving a leaner more profit-minded, but more impersonal, organization. The planning process is specialized by level of hierarchy, permitting more formal measurement of the process and the men who manage it.

7. Because return on investment in subunit ac-

tivities can be calculated, at least approximately, diversified firms as a group, try to allocate their resources to the most attractive opportunities. As noted above, this is not an easy process since it is not a direct matter to get the organization to generate plans and move people around. Opportunities are perceived narrowly and in personal and bureaucratic contexts.

8. Not surprisingly the problems that tax managements of the two sorts of firms turn out to be very different. Because top managements of the vertically integrated firms are so deeply involved in the planning process they are personally committed. The apparent result of this commitment is a narrower range of response to changes in the environment.

9. In the diversified firms, when the reward system is tied to measures such as return on investment, investment behavior is often responsive to changing economic circumstances. There are two cases to consider. In the loosely run conglomerate, where there are no hard, quantifiable standards for reward of managers, then investment behavior can be remarkably insensitive to changes in the economic situation. Instead, investment is largely a matter of momentum. More explicitly, size increases the ability to compete for funds and spend them. And dollars get allocated to business subunits roughly in proportion to their relative size. Because size is a function of previous growth, the implication is that divisions of a diversified firm grow at the rate of their sectors of the economy, plus or minus a variance potentially attributable to the quality of division management. Note that even in this instance where rewards are not tied to performance, there appears to be an inherent economic efficiency to the diversified firm. It will tend to allocate capital toward faster growing areas of the economy.

In the case of the tightly run firm, response may be much faster. If a top management can master the process of meta-management so that present performance is rewarded in a way that encourages future growth without "numbermanship," without a commitment to past glories; and without major bets on futures for which talents and resources are inadequate when measured against the market and competition; if all these conditions can be met, then the diversified corporation can represent a remarkably efficient instrument for moving resources toward the more rapidly growing sectors of the economy.

The business of making these adjustments takes time, but it is to precisely such problem that top managements of today's diversified firms devote most of their attention. Their focus is on

people, not things or numbers.⁶ Their principal tools are indirect; their time horizon is often a decade rather than one or two years; but they are managing the affairs of their company.

Contrast this picture with Galbraith's and the problem is clear:

The goals of the mature corporation will be a reflection of the goals of the *members of the techno-structure*. And the goals of society will tend to be those of the corporation. If, as we have seen to be the case, the members of the technostructure set high store by autonomy, and the assured minimum level of earnings by which this is secured, this will be the prime objective.

A main point of this paper is that Galbraith is right as long as managements attempt to manage in traditional ways. His "mature" corporation is in fact a large, traditional, vertically integrated, single-product company. Fortunately, I think, for our economy, the diversified firm, with the strength of its internal competition for resources, is increasingly the prevailing form of organization.

In Conclusion—Some Questions for Policy. It is appropriate in closing to note some of the policy problems posed by the highly diversified firm. First, while when measured over a decade, the diversified firm can be highly responsive to change, it can also be very hard to influence in the short run. The reason is to be found in precisely that indirection by which top management must proceed that was discussed above. Traditional economic policy is usually designed so that the economic attractiveness of business plans changes in one way or another; e.g., an investment credit changes the return on investment. The preceding discussion has suggested that a numerical analysis of plans is not critical to the approval process. The implication of that observation is that new policy tools may be needed. If, for example, the goal of an anti-inflationary policy is a slowing of the rate of private investment, some means must be found to induce the corporate officers to take pressure for current earnings and growth off their divisions. As long as division managers will rise and fall on relative profit and growth in the short run, they will put great pressure on their top management to find investment capital. The pressure of the inside management on top management is then likely to be very strong to keep investing, no matter how severe outside pressure may become.

Another critical problem arises in the area of

social policy. Let us take race relations as an example. The problems faced in its substantive detail by the operating units of the organization. In the diversified firm, these units are removed from the corporate top management by geography and by the formal organization and measurement systems. Unless pressure for earnings and return on investment are reduced, it is hard to see how or why division managements will pursue imaginative and aggressive—and typically costly—programs of racial integration.

An extreme example of the problem is provided by the behavior of the General Electric management leading up to the price fixing cases of 1962. The corporate management of G.E. required its executives to sign the so-called "directive 20.5" which explicitly forbade price fixing or any other violation of the antitrust laws. But a very severely managed system of reward and punishment that demanded yearly improvements in earnings, return, and market share, applied indiscriminately to all divisions, yielded a situation which was—at the very least—conducive to collusion in the oligopolistic and mature electric equipment markets.

Ralph Cordiner claimed that he never knew that price fixing was prevalent. No one has seriously disputed his claim in print. But consider the case were G.E. a vertically integrated single-product firm. Presumably instead of the division managers of G.E. who appeared in court, we would have seen the president. He would have played a major role in determining the company's pricing policy. Vertically integrated companies are likely to know what their presidents want. And logically enough, this seems to be true of integrated companies in the positive areas of social policy. The best records in the race relations area are those of single-product companies whose strong top managements are deeply involved in the business.

In short, the same forces in a diversified firm that tend to strip away economic fat and social tradition from the management of the enterprise tend also to strip away noneconomic aspects of all issues facing division managements, even those that are not remotely economic in character. The result is that while the planning process of the diversified firm may be highly efficient, there may be a tendency for them to be socially irresponsible.

A final problem is posed by the time horizon of division managers, and its relation to the kinds of investments division managements tend to make. One important aspect of innovation is that it tends to involve risk. The more inventive the development, usually, the more risk. It has been ar-

⁶ Solow could now argue that management is the scarce resource! And I will agree.

gued that diversification protects the technostucture against the costs of innovation and hence the diversified corporation can afford to be and is more innovative.

But, in fact, as we have seen, the risks to the division manager of a major innovation can be considerable if he is measured on short-run, year-to-year, earnings performance. The result is a tendency to avoid big risky bets, and the concomitant phenomenon that major new developments are, with few exceptions, made outside the major firms in the industry. Those exceptions tend to be single-product companies whose top managements are committed to true product leadership: Bell Laboratories, IBM, Xerox, and Polaroid. These are the top managements that can make major strategic moves for their whole company. Instead, the diversified companies give us a steady diet of small incremental change.

The conclusion to be drawn, I believe, is that planning in the American corporation has developed considerably in its sophistication but that it remains an enormously complex process subject to great uncertainty. The most important development has been the emergence of the highly diversified corporation as an agent of economic efficiency. On the other hand, the very effectiveness of the firm as an economic unit raises questions as to whether it will behave in a manner congruent with social objectives. As it becomes more tightly managed in terms of incremental economics, it also becomes harder to make larger, more global commitments that cannot be

economically justified against short-run measures. It is also questionable whether the effectiveness of business as an economic institution justifies the faith some have placed in its ability to lead a social revolution in either the city or the ghetto.

The problem with this analysis is that it leaves us with no clear-cut answers to the policy issues of our day. The aerospace companies are committed to their technology and product in the same way as other single-product firms. But nationalizing them does not help at all. Forcing the military-industrial specialists to diversify should increase their efficiency, but diversified firms to date have not proved as responsive to the country's social problems as may be desirable.

My own guess is that the government must take the leadership and break up the tasks of accomplishing a social revolution so that at least some of these tasks correspond to activities traditionally performed efficiently by private firms responding to short-run profit incentives. Carving up tasks in a systematic fashion may require a bureaucratic revolution, but my second guess is that such a revolution is in progress and will be peacefully accomplished.

But my guesses are beside the point of this paper, which is that, in order to achieve national or policy objectives by influencing the behavior of private firms, it is necessary to take into account the specifics of how they work. If we do, then we can harness the efficient but narrowly responsive power of private planning to the emerging agenda of public aims.

DISCUSSION

JOHN SHEAHAN: Every decade is entitled to at least one generous oversimplification. One of the more innocent of the 1950's was the belief that national economic plans would somehow generate more progressive economies. The 1960's sufficed, at least among development economists, for near-total disillusion. But Vernon's strategy for this panel rules out despair. He made sure to include someone with close knowledge of planning in France, where it has been associated with rapid growth and structural change. And someone able to make clear that diversified corporations able to unite local autonomy with planned reallocations among many fields are better able to cope with their environments and outpace the world around them. If planning to guide local units toward group goals works within the diversified firm, how can it be devoid of interest for a nation? And finally, in Carlos Díaz-Alejandro, he included one of those rare spirits able to combine intimate knowledge of Latin America with optimism about the possibilities of progress through improved planning.

While in Latin America, I was struck with the degree of interest expressed in French postwar planning. Much of the appeal came from a somewhat uncritical association of improvement in the French economy with the introduction of national planning. But the interest went well beyond this. It included admiration of the highly elaborate analytical apparatus developed in the later French plans, complete with ever more modern models. And, more importantly, with a belief that French planning worked through cooperation among producers and government officials, eliminating any need for the unpleasant prescription of conflicting interests expressed through competitive markets.

Cohen's contrast between the formal planning apparatus on the one hand and industry-level collaboration through the *économie concertée* on the other serves effectively to demolish any thought that it was the national plan itself that controlled events. He is surely correct in this respect. But, with a healthy determination to show some positive links between efforts and results, he may have overshot on the virtues of industry-level collaboration. In fact, his description sounds familiar: it is the "New Competition" of the 1920's all over again, with group cooperation replacing market pressures and a better world ensured by the intelligence and unselfish drive of the coordinators. This is clearly not a fatal prescription when the country concerned is blessed with a good number of progressive firms generating their own pressures for improvement. But it is clearly not a very good prescription for a closed economy in which both government advisers and their private partners like to make sure that enterprise financing is not put under excessive strains and that nobody really goes under.

Why did it work out reasonably well? Because it did not govern events. The companies that wanted to grow faster than the group went their own ways and sent amiable representatives to nod agreement in group discussions. And the industries that worked badly were undercut by external competition. When planning in consumer durables left the producers of

refrigerators with exceedingly high prices, consumers simply bought Italian refrigerators. Just as firms bought German machinery, and so on. Collaboration sometimes achieved useful changes but the controlling force was the fact that the long insulated French economy was both opened up to powerful external competitors and blessed with an exchange rate that made the better firms interested in exporting. Cohen's thesis, at least in this necessarily abbreviated version, puts the emphasis in the wrong place.

Bower's fascinating discussion of the problems of control in the diversified corporation is illuminating for national planning as well as for understanding of our potent but troublesome conglomerates. "The job of top management is to keep self-interest and corporate interest aligned." I cannot find an equally terse statement of the essential point in his consideration of the same problem at the national level, but his perceptive contrast between the conglomerate's operating efficiency and its lack of responsiveness to social goals makes the issues clear. The essential job of national economic policy is to align the corporation's self-interests with the goals of the society. It is not enough to spell out the targets and discuss them with corporate executives: policy has no meaning unless those taking daily decisions gain from pursuing the goals set by the society and lose when they go in the wrong directions. This is perhaps the most important criticism of Latin American plans in the paper by Carlos Díaz-Alejandro. How many of those plans in the 1960's included any explicit consideration of measures to align company profits with national goals?

Díaz-Alejandro goes in exactly the right direction: a search for ways in which planning could be used to improve market systems which were not very good in the first place. Planning in Latin America has not ruined previously efficient economies: it has simply been a disappointing irrelevance. The demand for more authority to enforce plans is essentially a plea to escape to a fantasy world in which it suffices to decide what must be done, in which all of Bower's well-defined problems of finding ways to induce people to harness their individual efforts toward consistent goals can somehow be wished away.

The potential gains of planning are still there. Many of them are on the side discussed so well by Díaz-Alejandro: that of improving efficiency. But it may be that the greatest contribution in the long run will come on the other side: that of more conscious integration of social goals with private consumer preferences. The economic dominance of the United States is compelling testimony to the power of a market system to generate efficiency on the supply side. The uses of this strength provide depressing examples of the consequences of failure to establish coherent goals. Latin American planning has so far done precious little good either on supply efficiency or redirection of effort in socially valuable directions. It is not likely that it will ever do as well as the United States market system with respect to supply efficiency. But it could, and certainly should, do much better at directing productive power toward creative ends.

JESSE W. MARKHAM: The striking commonality of the papers before us is the collection of evidence and arguments they offer to support the proposition that free enterprise economies of widely different types, even the individual free enterprises that comprise the larger share of these economies, plan; that planning in these economies is essentially functional rather than ideological; and that in broad purpose and design it is more a supplement to the market than a substitute for the market.

To be sure, the origins and components of the planning process differ among the national economies covered. In the United States, as Professor Eugene V. Rostow has painstakingly documented in his *Planning for Freedom*,¹ planning has been largely a matter of the expansion and evolution of public law; according to Professor Díaz-Alejandro, in Latin America planning arises out of the Hispanic-Latin tradition of the state's concern with economic regulation; and in their recently published *Industrial Planning in France*, John McArthur and Bruce Scott find *le plan français* to be distinctly French. However, in spite of these distinctive national characteristics, the planning process in free enterprise economies, except for such aberrations as agriculture, are ostensibly designed to comport with Díaz-Alejandro's prescription for planning the foreign sector in Latin America: it should aim at supplementing rather than replacing, or ignoring, the price system. Its priority target in Latin America should therefore be the maze of *ad hoc* import restrictions and tariffs accumulated over the last thirty years.

Cohen's diagnosis of planning in France may not at first blush appear to fit neatly into this pattern, but on close examination it is not inconsistent with it. As Cohen observes, planning in France is exclusively concerned with the large French corporation. The large French corporation operating in the French market rarely if ever confronts conditions generally associated with a state of effective competition. Hence, in France, too, planning begins where effecting regulation by market forces ends.

My faith in the price system may not be so unshaken, or unshakable, as to qualify me for membership in the so-called "Chicago School," but it is at least as strong as the average for the Greater Boston Area. This statement assumes of course that professor Galbraith is included in the average. For this reason I find myself in virtually complete agreement with most of the diagnosis and broad policy prescription in Díaz-Alejandro's paper. Where the market appears to function in the public interest, such as in the competitive textile industry, it would be a great waste of scarce planning resources to insist on planning. Such scarce resources are obviously more urgently needed, and hence their marginal product potentially much greater, in the areas of education and public enterprises, areas which most societies have subjected to the planning process, at least in the negative sense of having not entrusted them entirely to regulation by market forces.

¹ *Planning For Freedom: The Public Law of American Capitalism* (Yale Univ. Press, 1959).

There are awesome obstacles to this rational division of labor between government and the market on which Díaz-Alejandro touches, but does not dwell, perhaps because of their awesomeness. He rightly observes that the maze of *ad hoc* restrictions, riddled with internal conflicts and inconsistencies, grows out of uncoordinated powerful bureaucracies responsive to even more powerful special interests. The solution, if there is one, obviously lies deep in the political process, ultimately deep in the nature of man, but I should very much have welcomed a more systematic analysis of this issue than either Cohen or Díaz-Alejandro gave it. It is, after all, the central issue.

I find the terrain Professor Bower covers in his paper more familiar, although given the dissimilarity in topics he and Díaz-Alejandro develop surprisingly similar theses. To Díaz-Alejandro the imperfections and inadequacies of the market place provide the *raison d'être* for economic planning at the national and community level; to Bower these same imperfections and inadequacies, redefined as the discretionary power arising out of the absence of highly competitive market constraints, provide the firm with sufficient scope to plan. Hence, in the case of the textile industry the state need not plan; in the case of the individual textile firm it cannot plan.

I do not feel particularly moved to assess Bower's treatment of the Galbraith-Solow exchange, the profits maximizing assumption, or the games managers play and the rules by which they play them. I agree with his conclusion that Galbraith accords far more power to the firm to make its plans, including its mistakes, succeed than factual evidence supports, but this is simply to reiterate a modest quibble with Galbraith most students of the American economy have already voiced. However, I find one of Professor Bower's hypotheses concerning the differences in how large vertically integrated single-product market firms and large diversified corporations employ the planning process extremely fascinating. In contrast with the management of large single-product firms, large diversified firms react to a broad range of environmental changes since their multiple-profit center approach encourages them to be alert to investment opportunities wherever they may exist. The resulting internal competition for resources makes the diversified firm an especially appropriate institution for the allocation of resources to their most productive uses. But, according to Bower, these forces in the diversified firm that make their managers highly efficient tend also to make them less socially responsible—they are not simply profits maximizers, they are unbridled profits maximizers. A corollary to this is that large conglomerates tend to avoid highly risky activities because major innovations that go wrong can play havoc with a division's short-run earnings record—the scoreboard by which division managers are judged.

To go back to Bower's opening statement, I do not wish forever to disqualify myself for economics faculties by displaying even a modicum of knowledge about the firm, but students of industrial organization will quickly detect Bower's conglomerate as one bearing very little resemblance to the conglomerate image

emerging from the conventional wisdom. For two decades the large conglomerate has been berated as a slothful corporate creature insensitive to those commercial incentives promotive of efficient operations and profitability and as owing its survival largely to that mystical economic power that the corporate fates, assisted by Professor Corwin Edwards,² reserve exclusively for conglomerate enterprise. It derives its profits from all markets; hence it can pursue non-profits objectives in any market. Its one redeeming feature may be that, *ceteris paribus*, it will innovate more aggressively because the potential innovations that flow from a given research and development outlay will have a broader range of applicability within the firm; i.e., the risk attached to any given R and D outlay will be lower.

Bower has offered us a reverse image of how the large diversified firm functions. On the whole it is also a much more favorable interpretation of the conglomerate firm than it has traditionally received and for this reason is obviously not devoid of policy implications. The attack on conglomerates currently being mounted in Washington rests on the untested hypothesis that to curb their growth in size and number may provide some quantum of social benefits. The possible benefits justify the attack since the attack entails no social costs. Bower has developed the reverse image. Hopefully, factual analysis will someday disclose which is the true picture and which is the celluloid negative. According to recently released Federal Trade Commission data 181 of the 200 largest manufacturing corporations operate in at least ten different product markets. These large conglomerates account for over one-half of all corporate manufacturing assets and profits and for about two-fifths of total value added in manufacturing. The two conflicting interpretations of how firms accounting for so much of our economy actually function is obviously not a trivial matter.

ROBERT A. SOLO: Economic planning as practiced in Latin America, says Díaz-Alejandro, is a façade, a surface ritual with little if any relation to political practice and the flow of events. Hence it wastes the most precious of developmental resources: the time and energy of the trained, dedicated men with a place in the political establishment as planners.

Agreed. In the planning approach inaugurated at Punta del Este we said to our Latin American compeers, "If you want to ring the bell, then press the button, because that's the way it works in our house." Alas, if the electrical circuit between button and bell is broken? If the motor that rings the bell has never been installed? If there is not even a button to push?

Indeed, if I may anticipate, I am troubled that Díaz-Alejandro might also be devising a button pushing strategy rather than getting down to the nitty gritty of wiring circuits and installing motors.

Alternatively Díaz-Alejandro proposes programs for

Latin American planners that would (1) promote exports, (2) control imports, (3) achieve regional economic integration, and (4) stabilize and extend investment or credit inflows.

Díaz-Alejandro's concern is not that these things are not done but that they are not being done by planners with national goals and policies in mind. Rather they are tasks left to old-line bureaucrats in established ministries who produce a maze of *ad hoc* regulations in line with precedents and shaped by transitory pressures and vested interests. Only those who stand for and act in the name of a coherent national policy and strategy, Díaz-Alejandro would hold, can possibly withstand those pressures and offset the internecine struggle of ministries.

I am less sanguine than Díaz-Alejandro about the superior capacity of planners to perform these tasks. Planners easily preserve their pure souls, so long as they stay in the realm of irrelevance. Things change; men change when they are absorbed into the hierarchies of power and engage in the cross-currents of political choice.

There is more to it than that. Economic planners do what they have been trained to do. They apply what they have learned and what they have learned may be all that the discipline offers. If now what they do is useless, and we would make it useful, then we are after something quite different and infinitely more difficult than a reshuffling of tasks and a shift in personnel. What would be required is that the discipline itself be transformed, and perhaps that the socio-organizational frame in which the individual, *qua* planner or otherwise, operates be transformed also.

I am troubled by Díaz-Alejandro's sectioning off of "Planning in the Foreign Sector"—counting "domestic policies" simply as a parameter. That may be permissible where neoclassical theory adequately describes the relevant universe, with full employment, technological change, and productivity assumed out of sight. But not when economic development is at issue. What neoclassical theory assumed out of sight becomes then the crux of concern and foreign trade policy is significant only in relation to the inner transformation of society and reconstruction of the economy.

Withall, Díaz-Alejandro has an important message. Planning can mean many things. The meaning it should properly be given depends on the particulars of purpose and circumstance. For Latin American development, Díaz-Alejandro sees the planner neither as the lookout peering into the future through swirling mists nor as the navigator of the ship, neither as supercoordinator of decisions by public agencies and private firms in the French manner, nor as the engineer and controller of complex operations in the socialist mode but rather as an integral dedicated elite operating within and as part of the political mechanism for economic maneuver. He admonishes the planner to ride herd on the autonomous forces of the free market, spurring, offsetting, guiding, supplementing but always in order to make the most of and to get the most from the variety and creativity of these forces. In his conception, the goals of the development planner should be specific and limited—step

²See his "Conglomerate Bigness as a Source of Power," in *Business Concentration and Price Policy* (Princeton Univ. Press, 1955), pp. 331-59.

after step in an upward climb. And as each step is taken, the context and conditions for planning changes, and the planner must change his sights, redesign his strategies, and remobilize the forces at his disposal.

Díaz-Alejandro speaks to planning in a type of political organization and Bower to planning a type of corporate organization. Their conceptions are not dissimilar. Both see planning's role in the design and implementation of sets of strategies in pursuit of changing goals. For both the ultimate task of planning is to create a framework that can rationalize and exploit the self-interested drives of the decentralized operating entities: private businesses for the development planner; corporate subunits for the corporate planner.

These are also differences. Díaz-Alejandro addresses those who are designated as "planners" and who are officially engaged in "planning" for Latin American government. No group or function can be so specifically identified in corporate organizations. What Bower calls planning encompasses virtually everything that goes on in the company and might better be termed action taking, decision making, or simply choice within and by the firm.

Díaz-Alejandro's approach is normative. He evaluates the planning activity and urges its transformation. Bower's is analytic. He wants to explain what really goes on behind the corporate facade. He juxtaposes his conception of corporate choice against that of Galbraith and that of Solow and claims that his is the more realistic. But what is his conception? That is not easy to say. Let me try, however, to restate it. Modern corporate organization, he would hold, includes the activities and involves the interests of a very large number of persons and sheer numbers introduces qualitative changes into the relationship between participants. Problems are thereby created in communication within the enterprise and for the formulation, unification, and execution of policy. Modern corporate organization, moreover, is characterized by a wide dispersion of powers of decision making, of action taking, or of censure and veto. Why so? For one thing because the enterprise performs, not a few, but a very large number of functions, each of which may be carried to a point of great complexity and this complexity of functions proliferates specialists whose positions are rendered quasi-autonomous by the particularity of their competencies. The decentralization of decision making comes about, moreover, because competitive necessity requires that operations be decentralized and authority delegated in relation to the geographic spread of operations or to the diversity of products and/or sources. The orientations of those who participate in control are thus shaped by different functional experiences and responsibilities and are identified with a self-interest in different aspects of the operation. Hence, decision-making or decision-influencing groups or individuals will be characterized by disparities in outlook and goals and a complex operation will be required to control the complexities of operation. Besides the control of operations, corporate authority will be also directed to the maintenance and strengthening of organization *qua*

organization and the control powers will be used, moreover, to strengthen the position of and to benefit individuals or groups within the organization.

It is in relation to this interaction, counteraction, and multiple motivation that corporate behavior and choice must be understood and gauged.

If this succeeds in summarizing Bower's conception of the process of corporate choice, then I wholeheartedly agree with it. I should, since I have been quoting directly from an article I published a full decade ago.¹ I do not mean to suggest plagiarism. Quite the contrary. It would seem rather that Bower has never read or heard about anything that I (and presumably others) have written these many years on the subject with which his paper deals. He is satisfied to have assimilated the views of a few luminaries—which is unfortunate for one who sets up shop as an expert.

But if Bower's scholarly spectrum is narrow, he ventures to do what scholars have rarely done. Since organizational choice is not analogous to personal decision making and hence cannot be understood by an empathetic projection of a rational calculus in the neoclassical manner, Bower would turn rather to the direct observation of corporate behavior. He would be the novice who, rather than arguing as to the number of teeth in the horse's head by reference to authoritative texts, gets a horse and brings it in to show us.

Or so he claims. But, while I have heard him counting the molars, I have yet to see the horse. His paper is full of strong—indeed of absolute—statements about business choice, managerial behavior, corporate organization—but I cannot find the credentials for those statements. Why should I accept them as true? They are certainly not logically necessary nor observationally self-evident. Nor has the data been adduced to "prove" or substantiate them. Knowing something of what information is available, I doubt very much that it can be. Consider this, for example: that "major new developments are, with few exceptions, made outside the major firms in the industry." What does he consider a "major firm" and what "major new developments" in modern industry has he in mind? Computers? Transistors and semiconductors? Communication satellites? Catalytic engineering? Synthetic and plastic fibers? Talking pictures? Television? Color television? What?

Nor can I accept as self-evident the virtues for efficiency and progress that he so casually imputes to the "diversified firm." It is also possible to view the modern conglomerate as an organizational monstrosity with less technological rationale, but pasted together for the same reason as the public utility, holding company empires that collapsed so spectacularly in the late 1920's.

I have been critical, but I would say, nevertheless, that these are encouraging papers, controversial and challenging by ingenious young men who want to say things like they are. I hope these are harbingers of things to come.

¹ I have put together sentences from "Intra-Enterprise Conspiracy and the Theory of the Firm," *J. of Bus.* (Univ. of Chicago), Apr. 2, 1961, pp. 153–66.

NATIONAL COMMUNICATIONS POLICY

THE SINGLE-ENTITY PROPOSAL FOR INTERNATIONAL TELECOMMUNICATIONS*

By MERTON J. PECK
Yale University

When President Johnson established a Task Force on Communications on August 18, 1967, he specifically asked the group to "investigate whether the present division of ownership in our international communication facilities best serves our needs."¹ The Task Force concluded that the present organization should be replaced by a single entity to operate the long-distance portion of international telecommunications—a radical change from the present pattern.²

"Single entity" denotes the ownership by a single firm of cable and satellite facilities over which overseas telephone and record messages are transmitted. The Task Force's proposal would limit the single entity to the long-distance transmission function and prohibit the single entity from vertical integration into equipment manufacturing and affiliation with domestic carriers. Creating a single entity would bring together the overseas transmission functions of the Communications Satellite Corporation (Comsat), American Telephone and Telegraph (A.T.&T.), and the record carriers—Western Union International (W.U.I.), Radio Corporation of America (R.C.A.), and International Telephone and Telegraph (I.T.&T.). Comsat is the sole U.S. company participating in the ownership of satellites through the international consortium while A.T.&T. and the record carriers jointly own the cables. A.T.&T. alone installs and operates most such installations.³

* The author's more lengthy discussion of the topic is contained in *Essays in Honor of Edward Mason*, to be published by Houghton Mifflin, 1970. The author was a member of President Johnson's Task Force on Communications Policy and is much in debt for the opportunity to read the staff papers prepared for the Task Force, as well as for the comments of William Capron, of the Brookings Institution, Leland Johnson, of the RAND Corporation, Roger Noll, of the California Institute of Technology, and Harvey Levin, of Hofstra University.

¹ Global Communications System, Message from the President of the United States (80th Cong., 1st sess., Doc. No. 157, Aug. 18, 1967), p. 8.

² *Final Report: President's Task Force on Communications Policy* (submitted Dec. 7, 1968, released May 20, 1969). Henceforth cited as *Task Force Report* with page references to the mimeographed version.

³ The present organization of the U.S. interna-

The Task Force offers five reasons for recommending a single entity:⁴ (1) to promote system optimization and enable realization of the available economies of scale; (2) to further U.S. foreign policy objectives; (3) to resolve the anomalies of Comsat's role and function; (4) to help resolve the problems of the international record industry; and (5) to improve the prospects of effective governmental regulation.

This paper focuses on the first of these reasons. System optimization in telecommunications involves primarily the question of how to divide investment between satellites and cable. Both technologies are very capital intensive and fairly inflexible. Once the cables or satellites are in place, a good many other economic decisions, such as the division of output or current level of costs, are largely determined. Considered here are three questions: what is the economical combination of cable and satellite investment; why does the present organization of the industry and its regulation fall short of this economical combination; and in what ways would the single-entity proposal be an improvement?

I. The Economical Combination of Cable and Satellite Investment

International telephone calls and record messages travel over domestic telephone or Western Union lines to either satellite earth stations or cable heads for international transmission. For most traffic, cable and satellite transmission are almost perfect substitutes. The exceptions will be considered subsequently.

The Task Force requested the National Academy of Engineering to examine the relative costs of varying mixes of satellites and cable.⁵ The Academy's committee focused on the Atlantic Ba-

tional communications industry is described in greater detail in *Task Force Report*, Chap. 2, pp. 3-6.

⁴ *Task Force Report*, Chap. 2, pp. 27, 29, 31, 32 and 38, respectively.

⁵ *Report on Selected Topics in Telecommunications*, Committee on Telecommunications National Academy of Engineering (Final Report to the Department of Housing and Urban Development under Contract No. H952, Nov., 1968; revised Dec., 1968).

TABLE 1
COSTS FOR VARIOUS INVESTMENT STRATEGIES FOR
AN ANNUAL GROWTH IN DEMAND OF 10 PERCENT

Strategy	1985 Traffic via Satellite	Cumulated Annual Costs 1976-85 (\$ millions)
1. All incremental investment in satellites.....	85%	90
2. Incremental investment primarily satellites but with some cable.....	70	127
3. Incremental investment divided between satellites and cable.....	48	184

SOURCE: *Report on the Selected Topics in Telecommunications*, Committee on Telecommunications, National Academy of Engineering (final report to the Department of Housing and Urban Development under Contract H-952, Nov., 1968, revised, Dec., 1968), p. 109. Costs are those occurring each year which are associated with the planned investment for 1976 to 1985. These include annual maintenance, capital cost amortized in a straight-line fashion over the life of the investment, and an interest charge of 6 percent on the net investment.

sin, the area with the largest traffic volume, and on the 1975-85 time period, for investment already authorized by the FCC will meet the demand until 1975. Three strategies to meet a 10 percent annual growth in demand were considered: (1) all incremental investment in satellites; (2) a division between cable and satellite, with the emphasis on satellite; and (3) a second mixed strategy, with the emphasis on cable. As Table 1 indicates, the pure satellite is by far the cheapest. A fourth possibility—to invest only in cable—was not considered by the committee because some low traffic density locations could only be economically served by satellite. A pure cable strategy also would be considerably more expensive than any other strategy.

The pure satellite strategy is the cheapest because that is the lower cost technology for long distances and high-traffic densities. The pure strategy is optimal because of the very great economies of scale in both cable and satellite.

In addition, satellites have two qualitative advantages. First, their life is around five years as compared to twenty years for the cable, so that the rapid technological progress in communica-

tions can be incorporated more quickly into satellites. Second, satellites can collect signals over a wide area from a number of ground stations and so capitalize on their economies of scale. In contrast, cables are inherently point-to-point service, requiring extensive feed networks to achieve scale economies.

Cables have two advantages. They do not use scarce radio spectrum space; however, the scarcity cost of the satellite's use of spectrum can be minimized by locating ground stations away from metropolitan areas and bringing the signals to cities over telephone lines. Cables are also cheaper for short distances, though there is uncertainty as to the limits of their zone of advantage. Estimates range from 600 to 1,200 miles, depending in part on their traffic density, but, in any case, transoceanic traffic travels much longer distances than the maximum of this range.

If the only criterion were cost minimization and if the system were built *ab initio*, then a pure satellite system would be optimal. But there are some qualifications.

First, cables and satellites are not perfect substitutes for some traffic. Telephone conversations via satellite involve a time delay which becomes particularly noticeable when transmission is routed through two satellites. (A Tokyo-London call can pass through both an Atlantic and Pacific satellite.) The time delay is also a liability over even one satellite for high-speed data transmission. These two kinds of traffic require considerably less capacity than existing cables possess and thus create a need for maintaining existing cables but not for additional cable investment.

Second, there is an argument for a mixed cable/satellite system as insurance against service interruptions. Satellites have proved to be exceptionally reliable—considerably more so than cable. The hedge against satellite failure can be provided most cheaply by space satellites (both in orbit and on the ground) and by space ground station capacity.⁶

The cables are good insurance against only two kinds of situations: atmospheric conditions that

⁶ As this paper was being prepared, the first instance of a satellite failure occurred. In June, 1969, the Intelsat III Atlantic satellite became unusable because of an antenna system failure. Use of two inorbit satellites permitted restoring 90 percent of the nontelephone circuits. (One was the Early Bird—a satellite launched much earlier which was reactivated.) Television service was restored with greater difficulty by rerouting of Atlantic traffic through the Pacific satellites. (*Wall Street Journal*, July 15, 1969, p. 8.) Note, however, that cable played no role in recouping from the satellite failure.

blank out satellite transmission for short time periods and the jamming of satellite transmission by an unfriendly nation.

The Department of Defense is the sole customer who values highly protection against these two contingencies. The efficient way to provide this kind of insurance is to have the Defense Department subsidize the cable construction it considers essential rather than, as at present, simply to state a critical national security interest in a pending cable application to the FCC. DOD now pays only a share of the added cost through higher rates; most of the cost is borne by other users. If all the added cost were borne directly by the Department, the cost benefits to national security of cables could be evaluated relative to using resources to contribute to national security in other ways. This would remove the special levy on communication users for national security, which is comparable to having steel buyers pay for the steel used for national defense.

Other users also value the insurance against even short service interruptions provided by a mixed cable satellite system, but the premium this group would pay seems nominal. A pure satellite system represents a marked gain in reliability over the cable system existing prior to 1963. (Cables have been out for as long as two weeks; satellites, with one exception, less than an hour.) Again maintaining existing cables provides some insurance and, since their investment costs have already been incurred, there is little added cost to the system from this degree of diversity.

These considerations suggest that the optimal system would be the maintenance of existing cables with incremental investment concentrated in satellites, except where defense needs are overriding. Any added costs associated with new cable investment due to national security requirements should be borne by the Department of Defense.

II. Existing Regulatory Policy

The recent decisions of the FCC indicate their actual policy diverges quite sharply from this optimum. The FCC seems embarked on a policy of parallel expansion of cable and satellite. In 1966 the Commission authorized both additional cable and satellite service to Puerto Rico when either facility would have met the projected demand through 1973 and when authorizing both facilities roughly doubled the cost.⁷ Two years later the FCC authorized a fifth transatlantic cable.⁸ The

emphasis on the parallel expansion of both cable and satellite continues despite evidence that satellites are the lower-cost mode.

The pressures on the FCC for a policy of parallel expansion are very great. Neither the cable owners—A.T.&T. and the record carriers—nor Comsat wishes to be denied opportunities for expansion in the highly profitable and rapidly growing international telecommunications market. In addition, the Department of Defense has stated a national security interest in cable expansion in such instances as the fifth transatlantic cable, which will be the first laid to Spain and provides an alternative cable route to southern Europe. Subject to such pressures, the FCC apparently takes advantage of the inelastic demand by adopting policies that allow both groups to prosper despite the higher costs to consumers.

The behavior of the firms is also understandable. Comsat's main business is satellite communications and as the low-cost carrier they plan to grow with the expanding demand. The behavior of A.T.&T. and the record carriers is explainable largely by the quirks of regulatory policy. These carriers can own directly only cable and can obtain satellite circuits only by lease from Comsat. Present regulatory policy provides a powerful incentive for ownership as opposed to leasing since the carriers' allowed profits are determined by the capital investment ("the rate base"). Only ownership adds to the total profits allowed by regulation. This condition creates what has come to be called the "Averch-Johnson effect"; regulated firms tend to expand their investment over and above the capital intensity that would be achieved by competitive firms using the same technology and faced with the same factor prices.⁹ Thus, even a weak profit maximizing assumption would suggest that, as long as the allowed rate of return exceeds the incremental cost of capital, carriers will prefer owning cable to leasing satellite circuits.¹⁰

⁷Harvey Averch and Leland L. Johnson, "Behavior of the Firm under Regulatory Constraint," *A.E.R.*, Dec., 1968, pp. 1052-69.

⁸The FCC has attempted to meet the problem of the differential effect on the rate base of the carriers in using satellite and cable by granting the carriers part ownership of the satellite ground stations. In so doing, the FCC said that, without such an ownership "they [the carriers] would be faced with the prospect of an ever diminishing rate base, both in the absolute as well as relative sense, and would be driven to seek alternative means, not necessarily dictated by efficiency but by their need for survival." (*Ownership and Operation of Earth Stations*, 5 FCC, 2nd, 812, Dec. 8, 1966, at p. 815.) Yet in the choice between cable and satellite the rate

⁷*In the Matter of ITT Cable & Radio, Inc.-Puerto Rico et al.*, 5, FCC, 2nd, 823 (Dec. 7, 1965).

⁸*In the Matter of American Telephone and Telegraph et al.*, FCC 68, 589 (May 23, 1958).

There is an offsetting factor that favors the use of satellites by the carriers in the short run. The carriers originate (and terminate) all the traffic; Comsat does not deal with or charge customers directly. If the cost of leasing satellite circuits is less than owning cable, the carriers increase their total profits at any given rate by using such circuits for the overseas leg of the transmission. Yet the short-run increase in profits would be offset in the long run because regulation would reduce rates to reflect a smaller rate base. The short run and the long run are separated by the regulatory lag, the length of time between the realization of excess profit and the ordering of rate reductions. Whether this short-run effect dominates the long-run effect depends, of course, on their relative magnitudes and particularly on the duration of regulatory lag.¹¹ To date, such short-run gains have been insufficient to lead the carriers to forego cable investment.

III. *The Single Entity and the Economical Resolution of the Mix of Cable and Satellite Investment*

There are three ways of breaking away from the present policy of costly parallel expansion of cable and satellites: (1) promoting more competition between cable and satellites; (2) substantial changes in the present regulatory policies without a change in existing ownership structure; and (3) the single entity.¹²

The competitive alternative is superficially the most attractive to economists. Cost estimates of engineers are always subject to uncertainty. Competition presumably lets the market decide the question of the optimal mix of cable and satellite investment.

Yet this alternative requires two radical shifts

base addition from partial ground station ownership is considerably less than that in the ownership of cable.

¹¹ There are other considerations in the carriers' preference as between cable and satellite. Relative rates may not fully reflect the cost differences in the two modes. A.T.&T. and I.T.T. have cable manufacturing affiliates, and the carriers' control over daily operations is greater for cables. In addition, if the demand for international communications were sufficiently elastic, the loss in potential cable investments could be offset by increased use of domestic long-distance feeder lines to handle the increased traffic resulting from the use of lower-cost satellite circuits.

¹² The Task Force also considered another option: establishing competitive entities with each one owning both cable and satellites. The substantial economies of scale in each mode make this a highly uneconomical alternative. For the Task Force's discussion of this alternative and the three mentioned in the text, see *Task Force Report*, pp. 21-28.

in regulatory policy. First, the cable and satellite entities would need to be free to compete for the traffic of the few large users who could lease either cable or satellite circuits. Present FCC policy precludes this possibility, for it limits Comsat to leasing circuits only to the present carriers, "a carriers' carrier."¹³ Competition simply for the carrier's traffic alone—most of which originates with A.T.&T.—is too imperfect to be even minimally effective. Second, A.T.&T. and the record carriers must divest themselves of their cable holdings. Otherwise, these firms who originate most of the traffic would have an incentive to favor the use of cables.

Yet even with these measures, scale economies would still create a duopoly: a cable firm and a satellite firm. The two conceivably could engage in extremely cutthroat competition, since costs would be low relative to average costs. The final result might be a single firm. Traditional duopoly theory suggests that the more likely outcome would be a "live and let live" policy, based on a mutual recognition of interdependence. With price competition muted, the result would be market sharing and parallel expansion of both cables and satellites—a result very similar to the present.

The second alternative—more vigorous regulation—could well check the expansion of cable simply by denying new cable applications. The awkward problem of rate setting would remain. The optimum mix includes keeping existing cables in operation. (Depending on changes in technology, it may also entail the replacement of existing capacity.) If rates are set on the basis of the lower-cost satellite, the cable may no longer be able to cover its fixed costs—including a return on investment. If the cables need not be replaced, these costs can be regarded as sunk costs, and the losses of the cable owners would not jeopardize an economical outcome. These losses of the cable owners would be regarded as the risks of progress which investors must assume. But American regulatory policy exhibits a strong tendency to relieve investors in regulated industries of such risk, prompted, in part, because returns in regulated industries are held down to reflect the lower risks of a monopoly. The most economical strategy for achieving an optimal cable-satellite mix may have

¹³ *Authorized Users* Memorandum Opinion and Statement of Policy, 4 FCC, 2nd, 421 (July 20, 1966); *Authorized Users*, Memorandum Opinion and Order, 6 FCC, 2nd, 511 (Feb. 1, 1967); *ITT World Communications et al.*, Memorandum Opinion Order and Certificate, 6 FCC, 2nd, 511 (Feb. 1, 1967).

to face this tendency directly and devise ways to compensate cable investors. (This would also create a setting that would protect the option for cable replacement if that is desirable.)

The third alternative—the single entity—offers an opportunity to compensate cable owners by buying out the cable investment. The sunk costs of the cable would be borne by communications users generally and there would be no need to gear all international rates to the financial requirements of the cable system.

The single entity offers another advantage. As noted earlier, most communications users benefit at least somewhat from the insurance against service interruptions provided by maintaining existing cables, and, perhaps, depending on the future state of satellite reliability, on new cables when the existing ones require replacement. The single entity can internalize such benefits in the pricing and investment in a way that dividing the two communications modes into separate firms obviously cannot.

A single entity is not inherently a solution to the problem of rate base expansion. The single entity could still increase profits through a rate-base expansion from parallel cable and satellite investment or, indeed, any other form of excess capacity. The defect is inherent in the method of rate regulation, and the solutions lie in a radically different way of regulation or more vigorous regulatory control of investment. Still, regulators may be more willing to grant a single entity either a cable or satellite authorization than to deny either of two applicants its only way of expansion.

The advantages of the single entity are bought at the cost of foregoing rivalry between Comsat and the carriers for the long-distance international telecommunication function. While this rivalry generates excess capacity and added costs, it still appears to have been a significant factor in the very rapid technical progress in both cable and satellite technology. However, in the case of satellites, the rapid progress originated largely with the manufacturers of satellite equipment and the military and space programs. These sources of progress would still remain if the single entity could be precluded from vertical integration into equipment manufacture. Nevertheless, the possibility remains that, without competition, a single entity may be less vigorous in applying new technology.

The loss of competition could be partially off-

set by insuring that the single entity would be free of ownership and control by the domestic carriers. The carriers would no longer be proponents of cables but rather would have a strong interest in lower rates for the international traffic they originate. Under the Task Force proposal, the single entity would sell only transmission capacity, with the retail service function provided by A.T.&T., Western Union, and other firms. While countervailing power is often considered a weak reed in offsetting monopoly, A.T.&T.'s large size relative to an international carrier's may make the structure effective. The Task Force concluded: "And we suspect that the entity's principal carrier customer—A.T.&T., representing most users of international communications—would be quite able to ensure adequate performance from it."¹⁴

Another set of problems, while not arising from the single entity per se, are associated with moving to the single entity from the present structure. The present regulatory problem arises because the existing firms are unwilling to relinquish a growing, profitable market and are able to protect their role in international telecommunications. The same reluctance and political pressures must be overcome in the creation of the single entity. The same pressures would be applied to devise some way of sharing the market, with the sharing process not necessarily confined to international telecommunications. One could visualize some particular change in the domestic market—for example, the assignment of domestic satellite operations to the carriers—being given to compensate them for the loss of the international market and to buy their support for the single entity. But such an arrangement could well be too costly in terms of the opportunities foregone for improvement elsewhere. A single entity for international telecommunications is a distinct improvement, but at a stiff enough price it may no longer be worthwhile.

The transitional problem, however, applies equally to the other solutions. The basic problem facing any proposal for change is that present policies and structure are uneconomical for society but profitable for established firms. Yet the promise of lower communication costs creates a strong case for change, and among the various possibilities the single-entity proposal seems by far the most promising in reaching an economical mix of cable and satellite investment.

¹⁴ *Task Force Report*, Chap. 2, p. 39.

TECHNOLOGICAL ADVANCE AND MARKET STRUCTURE IN DOMESTIC TELECOMMUNICATIONS*

By LELAND L. JOHNSON
RAND Corporation

Domestic telecommunications supplied by regulated common carriers encompass a mixture of monopolistic and competitive elements, with the Bell System playing a dominant role. Consisting of the parent company, American Telephone and Telegraph, its twenty-two operating companies, Bell Laboratories, and Western Electric, the Bell System serves about 84 percent of the nation's telephone subscribers. The remaining 16 percent are served by a total of about 2,100 independents. Each company, whether Bell or independent, enjoys the exclusive right to supply public telephone service within the boundaries of its franchise and, in cooperation with the rest, provides its subscribers full access to Bell's nationwide network of microwave and cable facilities.¹ In addition to its local and message toll telephone service, Bell handles most of the transmission of television and radio signals from network program originating centers to their affiliated outlying broadcast stations. In competition with Western Union, Bell offers a variety of private-line voice, record, facsimile, and data services. Western Union, whose annual revenues amount to less than 3 percent of Bell's, holds a monopoly of public message telegraph service while competing with Bell in other areas. Both carriers offer a switched teletypewriter exchange service; at this writing, negotiations are under way for the acquisition of Bell's service (TWX) by Western Union. Within these categories interstate services are regulated by the Federal Communications Commission and intrastate generally by state commission.²

The industry is characterized by rapid technological advance interacting with marked changes in the level and composition of demand. Recent

years have seen an array of new terminal and switching equipment, along with continued development of terrestrial microwave and cable technology affording striking reductions in circuit-mile cost. Communications satellites potentially useful for domestic purposes are now under development. New and rapidly growing requirements for telecommunications are stemming especially from the burgeoning computer industry. Many groups have been clamoring to establish their own private communications system as a substitute for common carrier offerings. Demands have mounted to employ on common carrier networks so-called "foreign attachments"—terminals not owned and directly controlled by the carriers. The carriers have been beset by criticisms that they are unable or unwilling to accommodate adequately to new needs. Outside firms are proposing to compete directly.

One example of the potential threat to the traditional role of the carriers is the case of communications satellites. To the present day satellites have been employed only in the international field, largely for transoceanic service linking the continents. With continuing rapid technological advance, however, they hold promise also for domestic service—especially for distribution of television programming from network centers to broadcasting stations scattered about the country as a substitute for present-day terrestrial microwave. According to some studies, the cost of employing satellites over at least certain routes would fall substantially below Bell's current rates. Some fear that were domestic satellites owned and controlled by the established carriers, the savings afforded by the new technology would not be passed in full measure to the users. To perpetuate the virtual monopoly of television transmission would only ensure, according to this view, that the primary users of the satellite system would continue to pay relatively high rates.³

As another case in point, the concept of the specialized common carrier has gained force, partially in consequence of the growing needs of the

* The author gratefully acknowledges the useful comments of L. M. Chazen, R. E. Park, and L. M. Ross on an earlier draft.

¹ Commercial use of microwave radio frequencies for communication was introduced in the early post-war period. With radio relay towers spaced to provide a series of line-of-sight hops extending over hundreds or thousands of miles, microwave systems constitute the backbone of today's domestic long-distance network.

² The most recent extensive analysis of the industry is contained in the *Final Report, President's Task Force on Communications Policy*, Dec., 1968, Chap. 6; henceforth cited as the *Task Force Report*.

³ See the reply comments to FCC Docket 16495, Dec., 1966, particularly the submission of the Ford Foundation, the Communications Satellite Corporation (Comsat), and A.T.&T.; also the discussion in the *Task Force Report*, Chap. 5.

computer industry. The number of telecommunications-based computer systems rose nearly six-fold from 1963 to 1966. According to one filing before the FCC, nearly 50 percent of all computers are expected to employ telecommunications links by 1978. Many data users have questioned the ability of common carrier facilities, designed primarily to handle voice, to meet their future needs. According to one, "voice signals are usually transmitted in analog form, while the majority of data applications requires digital signaling. Data transmission also requires different bandwidths, error characteristics and holding times."⁴

Responding to these factors, the FCC has moved generally in the direction of expanding the play of competitive forces. In 1960 it liberalized rules enabling users to establish their own microwave systems as a substitute for leasing private lines from the existing carriers; and in 1966 it approved measures to permit virtually unlimited sharing by users of private systems. In 1968 it liberalized conditions under which users are permitted to employ foreign attachments on the public telephone network. In 1969 it authorized (by 4 to 3 vote of the Commissioners) operations of a specialized common carrier, Microwave Communications, Inc. (MCI), to compete with Bell and Western Union in providing private-line service between St. Louis and Chicago.⁵ At this writing several other prospective specialized carriers have filed for service, including one proposing a transcontinental system to link thirty-five metropolitan areas. While domestic communications satellites are not yet in service, the FCC is seriously considering the extent to which their eventual ownership and operation should be placed in the hands of major users, or in the hands of specialized common carriers, as an alternative to the traditional common carrier mold.

Most economists would applaud these moves. The existence of several firms in particular markets provides valuable yardsticks in judging performance. Competitive firms are more likely to experiment with new mixes of pricing and products in testing out new markets. (In this respect it is notable that MCI proposes to offer a lower-grade but also lower-priced voice service than now offered by Bell.) More generally, under a wide range of circumstances, one would expect

competitive pressure to contribute to both allocative and x -efficiency.⁶

At the same time, expansion of competitive pressure does pose complications that merit careful attention. Among them are terms under which established carriers should be permitted to compete with new entrants, effects of "cream skimming" by new entrants, problems of interconnection among competing carriers, and effects on the efficient use of radio spectrum. The purpose of this paper is to explore briefly these aspects.

Disparities between Rates and Costs

The specialized common carrier and the satellite cases both involve prospective entrants acting on the basis of cost comparisons between their own facilities and Bell's rates rather than Bell's incremental costs of supplying similar services. Wide variations exist between Bell's costs and rates for various services and routes. The rate for a telephone call is invariant with respect to whether it is carried over high-cost or low-cost facilities, over direct or circuitous routing. The rate for carrying a television signal is quoted in dollars per mile irrespective of whether the particular facilities employed embody the newest or the oldest in microwave technology. Moreover, common costs among services arise from the frequently required common inputs of indivisible plant and operating expense, with the proportion of particular services produced depending on the inclusion of other inputs specific to each service. For example, a microwave tower can simultaneously handle message toll telephone, private-line, and television traffic, depending on the inclusion of other inputs specific to each. Common cost comprises the difference between total cost and the separable or incremental costs identifiable with each of the services. The allocation of these common costs among services generally reflects to some degree their relative price elasticities of demand. In effect, services protected from direct competition bear a larger proportion of common costs relative to their output than do the more competitive ones.

The mere fact that a new entrant's rates for a particular route or for a particular service are lower than those of the established carrier does not indicate that the new entrant's costs are necessarily lower than the existing carrier's long-run incremental costs for comparable service. In order to discourage uneconomic entry, it is essential to permit the carriers to respond by adjusting their rates toward their own incremental costs. Existing rates must not be frozen to provide an

⁴A. J. Lipinski, *Digests of the Responses to the FCC Computer Inquiry*, Report No. 7379B-5, Feb., 1969, Stanford Research Inst., pp. 20, 53 and 57. This is one of seven SRI reports dealing with "regulatory and policy problems presented by the interdependence of computer and communications services and facilities," FCC Docket 16979.

⁵FCC 69-870, re *Microwave Communications, Inc.*, Docket 16509.

⁶See H. Leibenstein, "Allocative Efficiency vs. X-Efficiency," *A.B.R.*, June, 1966.

umbrella protecting uneconomic competitive activity.

In the past, the FCC has accepted grounds of "competitive necessity" in approving rate adjustments. But the notion of competitive rate reductions frequently faces resistance, since a larger proportion of common costs must then be carried by the less competitive services. That is, if a separate communications satellite system threatens to siphon away television traffic, a rate reduction by Bell to meet this competition would result in some of the common costs previously carried by television being transferred to other services; if private-line competition is faced on the St. Louis-Chicago route, then the common costs previously carried by Bell's private-line services over this route would be transferred to other routes or to other services. To many it seems unfair that some services protected from competitive entry should be forced to serve as residual contributors to cover whatever common costs the competitive services cannot carry. According to this view, it is not enough that the competitive service covers full long-run incremental costs. It should also cover its "fair" share of common costs—usually interpreted as "fully allocated" costs reflecting the relative time use of common cost facilities by the various services.

Application of fully allocated cost criteria is incompatible with liberalized freedom of entry. If entry is permitted, common costs will necessarily be shifted elsewhere—either by competitive pricing as described above or by umbrella pricing that forces a reduction in the output of the carrier's service exposed to the competitive pressure.⁷

At the same time, the danger exists of a carrier cutting prices to the point where revenues fall even below incremental cost in particular competitive markets, if it has protected revenues from other markets. Ordinarily, we would expect the unregulated profit maximizing firm not to price below incremental cost (except for short-term price cutting to drive out competitors) for to do so would reduce total profit. But if the firm is subject to a constraint on overall rate of return, it may do so even as a long-term proposition.⁸ It would appear essential to set a floor to rate reductions permitted to the carrier such that reve-

nues not fall below long-run incremental cost; otherwise, the carrier may succeed in driving out or discouraging new entrants even if its cost is higher for the competitive service.

Clearly, the task of estimating long-run incremental costs is beset by many complexities, including the dependence of incremental cost for a particular service over a particular route on the level and imposition of other services projected over this and other routes and the practical difficulties of delineating the long run.⁹ One of the severest challenges lying before the FCC in its movement toward encouraging competitive entry is an expanded staff capability and access to data permitting tolerably good estimates of long-run incremental costs.

The Problem of "Cream Skimming"

The objection is commonly made to entry into regulated industry that new firms are attracted to the more lucrative routes and services, while the established carrier, in meeting its "public responsibilities," is required to supply a variety of services in many areas that taken by themselves would not be profitable. The effect of cream skimming is to force a shifting of common costs to areas less subject to competition, such as public message toll telephone. For at least two reasons this argument has carried a good deal of weight in FCC deliberations. The first relates to distributional effects between large users and small users, where freer entry may redound especially to the benefit of the former. It was partially for this reason that one Federal Communications Commissioner, dissenting to the recent MCI decision, observed that the decision was "designed to cost the average American rate payer money to the immediate benefit of a few with special interests."¹⁰

The second relates to the alleged threat of cream skimming to the nationwide network. As two commissioners have declared in opposition to the MCI decision:

Our existing nationwide common carrier communications system depends on a nationwide or state cost averaging. No common carrier should be permitted to go into business, even on a limited basis, and select only the desired high capacity routes. You cannot say "let someone else serve these marginal areas." Rate averaging is an extremely important principle to the very existence of a nationwide common carrier system.¹¹

⁷ Problems of computing incremental costs are discussed by Wein, *op. cit.*, pp. 35-50.

⁸ FCC Docket 16509, *op. cit.*, "Dissenting Statement of Chairman Hyde," p. 1.

¹¹ *Ibid.*, "Dissenting Statement of Commissioner

¹ Circumstances under which fully allocated cost criteria may nevertheless be useful are discussed by H. H. Wein, FCC Docket 16258, FCC Staff Exhibit No. 50, July, 1968; a critical appraisal of the fully allocated cost standard is presented by W. J. Baumol in the same docket, Bell Exhibit 26, May, 1966.

² See the discussion of competitive price behavior in H. Averch and L. L. Johnson, "Behavior of the Firm Under Regulatory Constraint," *A.E.R.*, Dec., 1962.

This line of reasoning embodies a public goods component. With rate averaging, a nationwide network includes both "lucrative" routes (in which the incremental cost of the route is low relative to revenue) and "unprofitable" ones (in which incremental cost is equal to, or above, total revenue). Since users of the lucrative routes bear the common costs of the system, and perhaps provide a subsidy in addition, the unprofitable routes emerge to fill in the nationwide network. Because of possible needs in national or regional emergencies, so the argument goes, users of the lucrative routes benefit from the existence of the nationwide network even if they themselves do not employ the unprofitable routes. Hence, the fact that rates are high relative to costs on the lucrative routes merely reflects, in effect, payment by these users for the benefits they receive from the nationwide system.¹² However, since each user's individual contribution to the nationwide network is in any event very small, those on lucrative routes would individually be tempted to purchase from cream skimmers for the sake of cost savings to themselves. Only under a system of collective support involving protection of lucrative markets through barriers to entry, so it is said, can the maintenance and growth of the nationwide system be assured.

If concern about cream skimming is so widespread, one may ask why the FCC has nevertheless moved in a contrary direction. The answer lies, I would surmise, partially in the fact that the competitive sector we have been discussing runs to only about 15 percent of total Bell revenues, with the bulk of Bell's business still bound up in the public telephone service. As a practical matter, expansion of competitive pressure in this relatively small area may not force a substantial shift of common costs to the detriment of the small user or to the viability of the nationwide network. No less important is the fact that the pace of technological change has been sufficiently rapid in this industry to make all users better off over the decades, thereby making distributional concerns less pressing than otherwise would be

the case. Finally, given the existing large and viable nationwide network, the public goods argument is less compelling now than it was, say, fifty years ago.

The Problem of Interconnection

The degree to which competitive entry is attractive depends on the relationship between economies of scale and the relevant range of output. Thus, telephone service into a block of homes or business offices is regarded as natural monopoly in that the cable network required to provide any service to each point within the block is sufficient, at little additional cost, to provide all the service required within the block. Duplicative telephone lines and switching plant would encounter the same concern of intolerable waste as would the existence of parallel gas, water, and electrical power lines. In the case of long-distance transmission over high density routes, especially where users have a variety of special needs, the case for the franchised monopoly is less strong. Hence, the rapid growth of private microwave facilities, the emergence of specialized common carriers, and the prospect of communications satellites being owned and operated outside the common carrier mold.

However, in all these cases interconnection with existing carrier facilities is of crucial importance: the specialized common carrier will need access to carrier lines from his microwave terminal into his customers' downtown offices; and the television broadcaster might, in some cases, employ carrier supplied lines from its communications satellite ground station to his local station. Without a satisfactory interconnection agreement, these users would either be forced to uneconomic duplication in building their own facilities or, in the presence of prohibitive costs, they would simply be unable to operate. In the recent MCI decision, interconnection arrangements with the Bell System were left for later determination.

Clearly, restriction in interconnection can be a formidable barrier to entry. Access to carrier facilities at reasonable rates, in accordance with certain technical standards, is essential. The judgment of what constitutes "reasonable" rates will not be easy in view of the complications of common costs. In many cases, the interconnecting carrier facilities employed by outsiders will also be used for the established carrier's message telephone service, private-line, and other offerings. The basis on which common costs are to be allocated between the circuits handling these services and those handling interconnecting user traffic will remain a major issue before the FCC.

Robert Lee, joined by Commissioner Wadsworth," pp. 3-4.

¹² To be sure, users of the unprofitable routes also benefit from the nationwide system. But were their rates raised to reflect the benefit, the amount of communications service they demand would fall. Hence, the unprofitable routes would not be as fully developed, nor contribute as much to the nationwide network, as otherwise would be the case. Moreover, support from the lucrative routes through nationwide rate averaging is more appealing to regulators on grounds of "fairness," since identical rates are set for both classes of user.

Efficient Use of the Radio Spectrum

The established carriers have argued strenuously that growth of specialized common carriers and operation of communications satellites outside their own control would result in inefficient use of the radio spectrum. In accordance with present-day allocations of radio frequencies among various users, specialized common carriers and at least the first generation of communications satellites employed for domestic purposes will share the same microwave frequency bands with the established carriers. Many cases of interference may arise among these users. It is easy to imagine a specialized carrier's microwave relay transmission interfering with a nearby Bell microwave system; or a satellite ground transmitter interfering with terrestrial microwave facilities. Under current spectrum management practices, if a new user interferes with an existing user, the new user is usually forced to suspend or to modify his operations, even if the value to him of operating in the manner causing interference exceeds the loss suffered by the existing user as a consequence of the interference. No convenient mechanism exists under which new and existing users can arrange compensation between themselves in order to maximize jointly the value derived in employing a given spectrum space. Under these circumstances, a single carrier having complete control over a broad band of spectrum has an advantage over competitive firms in achieving more efficient spectrum use insofar as it can fully internalize all gains and losses among the particular uses to which the band is put.

The basic problem here, as Levin elaborates in his paper, is that no market system exists for allocating radio frequency space. Rather, rights to employ spectrum space are conferred by central authority to particular users or to classes of users with little leeway to permit transfer of rights from lower-valued to higher-valued uses. As diverse and conflicting uses of spectrum space continue to grow, it will become increasingly important to revamp management practices to facilitate efficient use among them. One may hope that the expanding competitive pressures themselves, making that much more painfully apparent the inadequacies of current management techniques, will assist in forcing the necessary reforms.

Concluding Remarks

In conclusion, two points are worth mentioning. First, in view of the expansion of competitive pressures, questions arise about the appropriate boundary between the traditional franchised monopoly and competitive portions of the industry. To what extent should entry be permitted in yet

other portions; or have we already gone too far? Given the present-day and easily projected technological basis of the industry, the boundaries might reasonably be drawn between the so-called "switched" and "unswitched" portions of the nationwide network along the lines suggested by the recent President's Task Force on Communications Policy. For a variety of reasons discussed in its report the public telephone service, requiring the switched network to provide access to any one of a multitude of points on demand, should continue to be offered on the basis of exclusive franchise. It is in the unswitched portion of the network, embracing largely private-line services, that seems the more attractive for competitive entry at this time.

This does not mean, of course, that competitive entry in private-line services is bound to be successful. Specialized carriers may thrive and multiply or they may go broke. Communications satellites may or may not eventually be merged into traditional common carrier operations. One of the more curious aspects of FCC deliberations is the extensive attention accorded questions of whether a prospective entrant (as in the MCI case) is financially sound and is prepared to offer services attractive to the public. Why such aspects cannot be satisfactorily resolved in the market place, as they are in the unregulated sector, is a continuing mystery.

Finally, it may strike some readers as curious that so much emphasis is placed here on competition, while on the international side serious attention is being given to the single-entity approach discussed by Peck. This difference in treatment arises largely from the enormous differences in traffic volume between the two areas. On both the international and domestic sides large economies of scale prevail. But in the former, even generous estimates of transatlantic traffic volume through the 1970's suggest that perhaps a couple of Intelsat IV satellites will suffice. On the domestic side, several satellites will likely be required to handle television traffic alone. International telephone traffic involving the U.S. might amount to as many as 96 million calls by 1980, compared with over 1.5 billion interstate calls placed in 1965 and about 6 billion estimated for 1980.¹³ Granted that making long-term projections is a very tricky business; still the continued gap in volume between the domestic and international sides may suffice to justify quite different market structures in the foreseeable future.

¹³ *A Survey of Telecommunications Technology*, Appendix A to Staff Paper 1 of the Task Force Report, pp. 18 and 49.

SPECTRUM ALLOCATION WITHOUT MARKET*

By HARVEY J. LEVIN
Hofstra University

The growing demand for spectrum—communication frequencies—has once more come to outstrip its supply. Electrical interference, congestion, and resultant spectrum scarcities threaten to make air, ground, and sea travel slower and less safe; construction, distribution, mining, and manufacturing more costly; data transmission and information processing networks harder to develop for wide use; and some potential advances in education, law enforcement, and national security difficult to realize.

This "silent crisis" seems due at least as much to deficiencies in our allocational practices as to any inherent characteristics of the radio spectrum resource. Today, rights are awarded gratis by the Federal Communications Commission and the President's Office of Telecommunications Management. With right-holders prohibited from selling any portion, they lack incentives to economize use today, to withhold current use if greater future value will result, or to transfer rights to others who may value them more highly. Large portions of spectrum are allocated to different services with no chance for interservice transfers despite the sometimes permissive policies that govern intraservice usage [24, p. 27]. Some consequences of this centralized nonprice system have been aptly described by the President's Task Force on Communications Policy:

Police and other public safety radio services in major metropolitan areas may be unable to obtain vital spectrum resources, while those resources allocated to other user categories go unused in the same area (e.g., frequencies reserved for forestry services were only recently made available to the New York City Police Department). The business community and the general public throughout the nation may be denied access to otherwise unused spectrum bands

simply because these bands are used for other services in a few metropolitan centers (e.g., land mobile services are unable to use spectrum allocated to television in areas where the allocations are unusable for TV). One class of users may be forced to adopt costly equipment modifications to meet growing demand, while another class, favored with an abundance of similar spectrum resources, may use them wastefully (e.g., private land mobile users have undertaken three major equipment revisions since 1950 to conserve spectrum, while certain other mobile services continue to use wider bandwidths than required by existing technology). New spectrum dependent services, irrespective of potential social or economic benefit, may be denied allocations or forced to adopt uneconomic design and operating practices to protect established services, without even the option to indemnify existing users against harmful interference (e.g., satellite services are forced to locate earth terminals in remote areas and to adopt sub-optimum system trade-offs and operating constraints to ensure absolute interference protection for microwave relay systems).¹

One farreaching "solution" long proposed by many economists would be to institute a full-fledged spectrum market unfettered by detailed regulation.² Users of spectrum would own rights to deliver signals of acceptable quality, to exclude others who would cause harmful interference, to use the frequencies for any and all purposes, and to transfer or sell them at will. Trading in rights would move spectrum from lower-value to higher-value uses and users until the value of any marginal unit of bandwidth is equal for all or until the cost of spectrum to any buyer equals its value to some next-best user. Users would also alter their mix of spectrum and nonspectrum inputs (wire modes, radio filters, storage space, personnel) until the value of spectrum saved just equals the cost of the substitutes used.³

* The main ideas for this article were first synthesized orally in an advisory briefing for the Telecommunications Committee of the National Academy of Engineering in Aug., 1968, and for members of the President's Office of Telecommunications Management. Prior research was conducted under a grant from Resources for the Future, where the formulation benefited from discussions at the RFF-Brookings Airlie House Conference on the Radio Spectrum, Sept., 1967, and from a Summary of Conference Proceedings by W. Kenneth Jones [19]. Special thanks are due to Merton J. Peck for valuable detailed comments on an interim draft, and to Leland L. Johnson for general suggestions.

¹ [12, Chap. 8, pp. 16-18]. For evidence on the magnitude and incidence of spectrum scarcities today, see *ibid.*, pp. 7-15, and [10, pp. 2, 7-18].

² See [2, p. 101]. This pioneering study contains the first detailed delineation of a full-fledged spectrum market. See also Rose [30, pp. 7-12]; G. E. TEMPO [9, pp. 50-70], more fully developed in De Vany *et al.* [35, pp. 1512-52], published after completion of my paper.

³ See, generally, [2, pp. 102-06], Coase [3, pp. 25-38], [4, pp. 40-45]. Rights must be defined in terms of signal outputs, not signal inputs as at present. See references [2, pp. 102-08] [24, pp. 30-32] [30, pp. 7-12] [35, pp. 1512-29].

Although this proposal has considerable conceptual merit and value as a point of departure for empirical work, the economist's exclusive concern with it has unfortunately polarized the issues. In the process we have overlooked important middle-range options between the polar extremes of a full-fledged market (which many high officials believe simply cannot work) and the current framework.⁴

In examining the present impasse, Section I will identify certain economic obstacles to the formation of a complete spectrum market via private contract and common law. However, these obstacles have little to do with the viability of several modified market-type arrangements which meet the critical noneconomic objections the government still has to almost any market system (Section II). Therefore, Section III will offer three ways to inject price incentives more effectively into our present centralized nonprice system. The conclusion is that now is the time to pay far more attention to the mechanics of such promising middle-range options as user charges on occupied bandwidth, auctions within zonal constraints, and shadow demand prices for spectrum without markets. In this, my position gibes generally with that of the Communications Task Force which, in rejecting a full-fledged market at this time, endorsed a more "eclectic approach."⁵ Because these options are more "salable" than a complete market in practical, strategic, and political terms, the burden of economic proof they must meet as any preferred way to move us off dead center is far less onerous.

I. *Economic Obstacles to a Full-fledged Market for Spectrum*

Several important cost externalities impede the efficient definition, enforcement, and reconstitution of rights in spectrum and hence the creation of a viable market via private contract and common law. One problem relates to the variability in signal outputs due to natural phenomena beyond the user's control which produce extensive and unpredictable patterns of interference. The best examples are those of nighttime "skywave" in the AM broadcast band and of the outages and aber-

rations of high frequency radio transmission during the periodic sunspot cycle.⁶ A second, closely related externality arises from the general unconfinedness of radiation irrespective of spectral region or transmission mode. Here rights to radiate on designated frequencies in designated areas often have unpredictable or unauthorized effects on other frequencies in other areas. This is the problem of spurious emissions generally and of harmonic and adjacent channel interference in particular. A third externality relates to intermodulation effects where several different services operate simultaneously on different frequencies but in the same limited physical area (mountain tops, urban building roofs, naval vessels). Unlike the first two externalities, neither C nor B alone would harm A, but in C's presence B does harm A, through no fault of his own, while C harms neither. A further complication follows from the fact that interference by B of A's reception (in C's presence) may be due more to the low quality of A's receiver than to the power of B's transmitter.⁷

These three cost externalities and the uncertainties associated with them raise questions as to whether single-valued radiation rights can be readily defined, measured, and enforced. Without these prior conditions, of course, rights cannot be bought, sold, or reconstituted efficiently in a bona fide market. But property rights will emerge to internalize external costs and benefits and hence facilitate efficient market transactions only where the cost of internalization falls short of the gains that result [7, pp. 348-49]. Where the interactions of the right-holders immediately involved affect many others besides or are in turn affected by them, private negotiations to reconstitute and police the rights are normally considered inefficient and inappropriate [7, p. 357] [3, pp. 29-30]. This is true even in the absence of a more efficient government mechanism to do so (in which case efficiency requires that externalities simply be ignored) [5, pp. 15-18, 26-28, 41-44] [8, p. 12].

As a practical matter, the reconstitutability of rights to radiate will depend on the number of parties to representative transactions. Enforcement and transfer costs will be lower the fewer the parties who must negotiate and whose subse-

⁴ One outstanding recent exception appears in the *Task Force Report*, published well after completion of my own research. See reference [12, Chap. 8, pp. 28-38, 63-64]. The Task Force's conclusions differ on major counts from those of the searching outside study it had commissioned into the economics of spectrum management. See [9], elaborated in [35].

⁵ See [12, Chap. 1, p. 23; Chap. 8, pp. 30-33]. It should be noted that my analysis was undertaken independent of the Task Force inquiry and without knowledge of its staff papers or *Final Report*.

⁶ See generally, Joint Technical Advisory Committee [17, pp. 112-27, 218-19 (AM skywave); 127-42, 220-26 (HF skywave)] [16, pp. 33-36, 53-55, 156-57 (AM skywave), 65-91, 159-63 (HF skywave)].

⁷ A thorough statement of the engineering aspects appears in Joint Technical Advisory Committee [18, Sup. 6].

quent activities must be policed, whereas markets will be far less efficient where the number of transacting parties is very large and the withdrawal of any single participant can prevent a satisfactory agreement [5, pp. 15–18] [8, pp. 347–50]. The latter facts will impede the emergence of a market through private contract, but rights could conceivably be defined by the government first and then sold outright or leased for renewable periods to the highest qualified bidder.

Proponents of the full-fledged spectrum market assert that once rights are defined as signal outputs and users left free to combine transmission inputs at will, optimal degrees of interference will result. The definition of such output rights is of course theoretically conceivable. But their operational relevance cannot be assumed out of hand nor demonstrated by a priori deductive theorizing alone. Explicit empirical documentation of a sort not yet forthcoming is absolutely essential before one can expect public administrators to trade off the certainty of an ongoing system however imperfect for the paper promises of a hypothetical ideal.

One critical problem is that output rights can be defined only in probabilistic terms. If all the rights in adjacent geographic areas or on adjacent frequencies are used simultaneously; e.g., no single right-holder can predict the interference he will create (or suffer)—unless he knows the transmission inputs of all other right-holders. Without some minimum degree of certainty, on the other hand, market transactions in spectrum rights are inconceivable. Yet the minimum degree of input specification needed to insure this modicum of certainty may well preclude transfers of spectrum across present band or service lines. Specifically, “insufficient” input specification may reduce the value of the rights (via uncertainty) to the point where even modest exchange or enforcement costs preclude a viable market, whereas “too much” input specification will so increase the number of parties to representative transactions, that the cost of internalizing the externalities associated with spectrum utilization could well eliminate any net economic gain derived from exchanging even valuable radiation rights. These serious problems (and there are others) cannot be argued away on any abstract theoretical basis. Systematic empirical inquiry is of critical importance.

Meanwhile, individual negotiation and enforcement are widely agreed to be impractical in the case of man-made noise (due, e.g., to medical, industrial, or scientific radiation, to auto ignition switches, or street cars). Here the number of in-

dependent sources of radiation is very large, their detection costs high, and the number of people transmitting and receiving the noise excessive. Special administrative regulation is therefore needed to set standards in regard to receiver quality, hardware design, transmitter shielding, wiring, etc.⁸ Would individual negotiation and enforcement be very much more manageable in regard to major communications or locational uses of the spectrum?

The answer is not black and white. On one hand, the cost of internalizing the externalities doubtlessly has declined over time in response to significant technical advances in directing and confining AM broadcast transmissions and predicting the propagation conditions of HF radio [34, Chaps. 6–7] [28, pp. 41–51] [17, pp. 142–58]. Even more dramatic improvements are implicit in the advent of line-of-sight technology as such. In the VHF, UHF, and microwave regions generally our ability to predict the precise three-dimensional space any spectrum user will occupy and to define radiation rights accordingly are doubtlessly superior to comparable predictions in the older, lower-spectral regions. This probably deserves all the fanfare it has recently received, for there is no question that propagation vagaries have significantly colored our whole attitude towards the spectrum and its management.

On the other hand, we must not exaggerate the significance of technological advance. Both HF and AM broadcast transmission continue to pose serious interference problems. To cope with these in the AM band, regional agreements are periodically negotiated among numerous North American nations [26, p. 16]. Even within the U.S., difficult AM interference patterns show up in complex hearings where numerous parties from far-flung communities may protest some particular licensee's application for greater signal power or a longer broadcast day. The detail in which the FCC must now specify transmission hardware requirements further underlines the kind of agreements that would otherwise be needed between the individual users affected [26, pp. 9–13] [13]. In the HF band, far-flung global interference patterns also persist [26, pp. 23–27]. To manage these, complex negotiations must still be conducted [32].

Even with line-of-sight radiation, special interference problems may render the definition, enforcement, and transfer rights hazardous and costly. Without the proper (and possibly costly) hardware adjustments, e.g., a TV broadcast satel-

⁸ Coase, Meckling, Minasian [2, pp. 128–33]. See generally JTAC [18, Sup. 9, pp. 2–6, 19–25].

lite could conceivably interfere with numerous aeronautical, marine, or land mobile systems and especially with conventional broadcast and microwave transmissions. And it could do so on the same, adjacent, or harmonic channels over one-third the earth's surface.⁹ Individual negotiation of the necessary transmission standards would obviously be unwieldy. Radar offers a still better example of the hazards of overgeneralizing about the new line-of-sight technology.

Spurious emissions on adjacent and harmonic channels are simply much harder to contain in the radar bands than elsewhere in the microwave region or than in the FM and TV broadcast bands. Unless properly engineered, a radar facility overlooking a sizable city could potentially interfere with all radio services operating there, across the whole spectrum. The individual transactions needed to negotiate an optimal level of interference would be far too numerous, cumbersome, and costly to handle on any market basis. As with medical, scientific, and industrial radiation, an administrative designation of transmission standards and design objectives is essential. So, too, is the concentration of all transmitters in a few wide exclusive bands.¹⁰

In the case of intermodulation on a mountain top or urban building roof,¹¹ finally, it is at least conceivable that rights could be defined, liability for damages assessed, and individual negotiations conducted to unravel the situation. But the practicality of such individual negotiation is another matter. One way out might be to lease or sell outright the whole mountain top to a single licensee, leaving him free to sublet physical sites and frequencies to different users. The licensee would presumably preengineer the whole system (receivers plus transmitters), spacing and charging each user so as to maximize total rental income.¹² But the question is whether such unitary management would emerge efficiently, if at all, through individual negotiation alone (via an interfirm agreement, merger, or otherwise).¹³ Or whether any

such solution would initially require explicit federal designation of appropriate spatial and temporal tenure arrangements, albeit with subsequent lease/sale to the highest bidder.

In short, cost externalities will pose a potent economic obstacle to the emergence of any efficient market through private contract and common law, and this is true notwithstanding the technical advances just described.

Benefit externalities pose a further impediment to market efficiency, albeit one easier to handle through appropriate regulation. The broadcaster's access to virtually all resource inputs, e.g., including spectrum, will now be nonoptimal simply because advertiser-supported broadcast television cannot collect payments from all who view its programs. The institution of pay TV would presumably eliminate many such externalities [25, pp. 77-80] [31, pp. 81-83]. Under the present system of company-subsidized blue-ribbon programming, moreover, external public service grants-in-aid might help handle still others. But meanwhile, the obstacle posed for an efficient spectrum market seems quite clear. And this is true also of scientific, educational, and governmental spectrum users, which are equally unable to collect direct payments for all the services they render.¹⁴

II. Noneconomic Obstacles to Full-fledged Spectrum Markets

Economic obstacles to one side, the major opposition to a full-fledged spectrum market or even to market-type systems far short of that arises from a number of subjective judgments of the frequency management establishment. These are for the most grounded on untested assumptions regarding the consistency of such markets with prevailing regulatory doctrine. These noneconomic objections may indeed be deemed so persuasive today just because policy-makers incorrectly equate market incentives with a full-fledged market and fail thereby to distinguish the latter from management systems that utilize shadow prices, user charges, or stratified auctions.¹⁵

One complaint is that the U.S. image, already tarnished by neo-Marxist stereotypes, would be

⁹ See generally, Haviland [15, pp. 8-14] [14], *passim*. Cost implications of alternative systems options are explored in Communications Satellite Corporation [6], *passim*.

¹⁰ In all fairness to the market proponents, they have largely limited their most rigorous analysis to the region between 50 and 1000 Mc/s [35, p. 1502], thus abstracting from those intractable international and technical problems of AM and HF radio interference, and from the special problems posed by radar. Yet about one-half of the economic value product associated with the radio spectrum in 1962 was generated outside the 50-1000 Mc/s region.

¹¹ See textual discussion associated with footnote 7.

¹² See generally, JTAC [18, Sup. 6].

¹³ Compare with Demsetz [7, pp. 354-58].

¹⁴ Actually there is no more reason to exempt public users from market contests for spectrum than for their other inputs (now normally bought in relevant markets). And the public goods element in governmental or educational radio services poses no greater impediment to market efficiency than that in governmental or educational activities more generally.

¹⁵ Although too often blurred in the literature, these important distinctions were made explicit and clarified in Jones's excellent paper [19, pp. 85-105].

further impaired by any international spectrum market which prices out the less affluent nations [27]. We have apparently discounted the possibility that a redefinition of rights coterminous with national boundaries could enable the poorer nations to share in spectrum value by leasing theirs to others, without them (the owners) using it directly at all [2, pp. 207-10.]. Or that such a leasing system could hardly be interpreted as U.S. aggrandizement. Or that, if phased in gradually over time, it would raise far fewer questions of equity and political acceptability than any scheme to vest ownership (without compensation to the government) in those who currently use it. Nor does it in any case follow that a domestic market is impractical just because a bona fide international market may pose special problems.

A second fear is that even a strictly domestic market may be unworkable because some domestic radio services, though confinable, require foreign coordination and cooperation because they operate near our territorial borders (land mobile, FM or TV broadcast), whereas others do so because they are not easily confinable (AM broadcast). Still other services require close international coordination by their very character (international fixed, mobile, broadcast, space) [26, pp. 1-2, 6].

Perhaps the only really substantive international constraint is subsumed under a third factor: the rule of common use of common frequencies. Frequencies allocated to a strictly domestic service one day—and thereby ostensibly subject for a market—may subsequently be needed for a new international service. In the case of space satellites and radio astronomy, e.g., ITU members had to clear certain domestic bands to accommodate a new international service. Such accommodation was doubtlessly rendered more difficult and costly just because the several nations had failed to synchronize their spectral location of various domestic services during the wartime hiatus of ITU meetings, 1938-47. Whether a domestic U.S. market is conceivable nonetheless really depends on the frequency with which such far-reaching technical changes are likely to occur and the adequacy of eminent domain in handling them.

A domestic market is opposed, fourth, as likely to favor a few big wealthy users in private commercial services at the expense of public, educational or scientific users [2, pp. 165-67] [19, pp. 90-92]. Public accountability and budgetary review would certainly become more rigorous in a spectrum market. As a consequence, the allocators' discretion would presumably decline, as would the public users' discretion to combine

spectrum with other inputs at will. But the latter must now buy most of their nonspectrum inputs anyway, so why not spectrum too? In theory, spectrum would go not to those best able to buy it (the wealthiest), but to those for whom it has greatest relative value [3, pp. 19-21]. Indeed even today, the wealthier interests in broadcasting have gotten more and more assignments anyway, whereas in a new regulated system with prices special safeguards could prevent undesirable changes in industry structure [23, pp. 29-31].

Administrative discretion would be narrowed as feared, fifth, wherever allocators encounter adjudicable property rights. Yet from this it need not follow that the property rights themselves produce these constraints so much as the absence of markets where they can be traded. Nor are the regulatory priorities threatened by trading necessarily worth safeguarding anyway, where a cumbersome, inefficient, centralized system is the only approach.

Still another reason for the regulator's resistance to any market system is a fear of triggering thoroughgoing opposition by users who had hitherto gotten their spectrum for nothing and would therefore fight bitterly against charges of any sort being levied. These fears are understandable. But they must not obfuscate the possibility of "phasing in" a market-type system gradually or the varying ease with which different systems options could in fact be instituted.

Short of the most extreme abuse, finally, the momentum of an ongoing framework goes far to explain the regulator's widespread reluctance to consider any market system seriously. Numerous skills would be obsolesced, deeply ingrained habits of thought scotched, and the certainty derived from having immediate access for priority uses lost. Whenever an ongoing system, however imperfect, is compared to any hypothetical alternative, however persuasive, the existing system has all the advantages of actuality. Given the far-reaching institutional changes required this is even truer here than generally. It is true all the more because partial patchwork remedies, though far from ideal, can still be instituted to improve the situation.

III. *Towards a Regulated Market-type System with Prices*

In view of the above discussion, there seems much to commend something less radical than a full-fledged market. Properly conceived, a number of middle-range options might help us overcome the major obstacles (high enforcement and transfer costs) in a fashion that also sidesteps impor-

tant sources of political opposition. General administrative discretion would be limited and constrained, not eliminated. It would be forced to take into account more of the relevant facts before implementing any specific regulatory priority. Public users would have to weigh the cost of spectrum at least as carefully as they now do of nonspectrum inputs.

Ranked inversely to their nearness to the current nonprice management framework, the major options and suboptions are listed in Table 1. There seems little question that, if we had it all to do over again, a system of freely-transferable rights that worked would be by far the best from a strictly economic viewpoint. However, such a system seems farthest from current real-world practicalities and the option least likely to gain national acceptance. Nor does the specific version discussed most widely of late—based on interband auctions, transferable rights, and their eventual sale outright (A-2, 3)—appear viable in any case.

On the other hand, administrative techniques like frequency clearance and secondary rights (E) are in their own right important evidence that spectrum management has already begun to recognize the importance of injecting economic constraints into the present management system. More work can and should be done on this approach and valuable advances can probably be made. But if we are serious about articulating economic factors in spectrum management today, we cannot stop with this administrative attempt to find "equivalents" for the pricing system. Just as we are nowhere near ready for a full-fledged market for spectrum, so we are already well beyond the point where we can stop with a purely administrative approach.

What we really need is a more eclectic strategy that meets at least three requirements. It must be sufficiently close to the current framework to be relevant. It must contain enough economic merit to make it attractive on that score as an important step forward. And it must be amenable to being phased in gradually in ways that reduce the likelihood of what could be lethal opposition by those with "line authority" and real-world responsibilities. The present managers and users of spectrum will understandably feel that an ongoing reality, however imperfect, is decisively preferable to the paper promises of any hypothetical ideal, however persuasive in theory. Hence the allocational innovator must bear a heavy burden of proof.

The approach that comes closest to meeting the above requirements would I think be a sequential approach grounded in options D, C and B-2, 3;

TABLE 1
ALTERNATIVES TO THE CURRENT SYSTEM
OF SPECTRUM ALLOCATION

-
- | | |
|---|--|
| A. Freely-transferable Rights. | 1. Rights created in the courts under tort law, through inclusion and exclusion, with spectrum bought and sold outright, like land. |
| | 2. Federal designation of rights, leaving them freely-transferable after their subsequent sale outright. |
| | 3. Federal designation of rights, periodic competitive leasing for limited periods, with lessees substantially free to transfer at will. |
| B. Auctions of Federally-designated Radiation Rights. | 1. Interband contests to determine reallocation as between different services, in addition to intraband contests limited to like users within the same service. |
| | 2. Interband contests to ration grants among like users, within different services competing for the same spectrum, with managers free to utilize the resultant values in further reallocation of spectrum between the two services. |
| | 3. Intraband contests within a single service to ration rights there, with results used to set user charges elsewhere too. |
| C. User Charges. | Applied on occupied three-dimensional spectrum as measured by some index of physical usage. |
| | 1. Per unit rates derived from intraband auction values. |
| | 2. Per unit rates derived from estimated shadow prices. |
| | 3. Per unit rates set at some arbitrary flat dollar rate. |
| D. Shadow Prices. | Derived from maximum sums that current spectrum users and systems designers would be willing to pay rather than do without some small amount of spectrum. |
| E. New Administrative Techniques. | Greater role for frequency clearance,* secondary rights,† and a heavier burden of proof on spectrum managers where they deliberately override economic considerations. |
-

* The requirement that federal government users secure prior authorization to use particular frequencies (from the Director of Telecommunications Management) before disbursing funds to develop or build any communications system.

† Rights to share or borrow frequencies contingent on noninterference with rights of the primary user.

viz., shadow prices, user charges, and intraband auction values.¹⁸

¹⁸ The *Task Force Report* emphasizes the second element [12, Chap. 8, pp. 34-37], with additional proposals to "permit greater transferability of licenses among legitimate spectrum users within broad service classifications" (*ibid.*, p. 37). Shadow pricing

A. Market Simulation and Shadow Prices. By shadow prices I refer here to the maximum sums users of spectrum would be willing to pay rather than do without some marginal amount of bandwidth. These are at best only very tentative and imperfect approximations of the true shadow prices, whether the "commercial" prices which would result from unregulated (perfect) competition, or the "ideal" prices that would measure marginal social opportunity cost for all users, public and private [1, pp. 240-42]. A distinguishing characteristic of shadow prices is their estimation of the value of spectrum without any actual buying, selling, or leasing of frequencies. Once calculated, the spectrum manager has three options: to allocate spectrum accordingly; to override the "economic verdict" for cause; to translate the shadow price into a set of charges on occupied bandwidth.

In this last regard, a tax on user A equal to the extra costs incurred without some marginal unit of spectrum would reduce A's use depending on his elasticity of derived demand for spectrum [1, pp. 252-53]. If A were a big commercial user, the "commercial" shadow price for spectrum would fall accordingly. By the same token, the prevailing shadow price at the outset (as inferred empirically from the shadow demand price for spectrum) could give a rough "feel" of the spectrum value in question. This could help further in setting a proper unit rental charge to apply against the total volume of occupied spectrum. On the other hand, rental rates could also be set on an arbitrary flat basis or according to auction values derived in selected services.

Even assuming a systematic governmental estimation effort, little more than the crudest shadow prices are likely to be available for a long time [19, p. 101]. Furthermore, shadow price values will presumably reflect economic values in the current rather than the potential uses of spectrum. Yet in view of the arbitrary manner in which allocational decisions are now made, even crude measurements are better than none. Finally, it should be possible to look to the actual trade-offs implicit in the existing stock of communications hardware for some tentative feel of the actual values involved [22, pp. 481-84]. This would obviate any excessive dependence on the costs the hardware designer says he would incur without x -units of spectrum.

Basically, what we want to know is how net

value output (net of exclusionary cost) will vary as additional bandwidth is allocated to A and B (or reallocated as between A and B). Therefore, even a crude estimate of the value of spectrum to alternative users must consider the extra value of output generated in either service after an allocation of additional bandwidth, the extra costs incurred by either service, and the amount of extra bandwidth allocated.

Strictly speaking, the relative value of some small unit of bandwidth to alternative services can be estimated only if we take into account all three factors. For in many real-world cases the contending services each require different amounts of extra bandwidth and hence the extra net value output cannot be directly compared. Where the probable value output and exclusionary costs both remain constant, however, the service that uses less bandwidth should clearly be favored on economic grounds. Where the services use comparable bandwidth, the one that generates greater net output (over and above the extra costs due to exclusion) should be favored.

Finally, if the net economic value output is constant or even favors A over B, the spectrum managers may still conceivably favor B over A where, e.g., the regulatory priority associated with B's output is demonstrably higher than A's (e.g., safety, security, or education versus industrial usage). But even here, applicants' claims about "merit" uses will obviously be subjected to much closer scrutiny.

B. Rental Charges. Once the shadow demand prices of spectrum for alternative services are calculated even crudely, the resultant values could be used to help design an appropriate rental charge schedule. The same is true also of values derived from auctions. Using occupied spectrum as the multiplicand, per unit charges could actually be set in at least three ways:¹⁷ at some arbitrary flat rate; at rates related to shadow demand prices for spectrum, as derived from the costs imposed by exclusion; and at rates that reflect auction prices on an *ad hoc* basis; or rates geared to cover the cost incurred by a private concessionaire in acquiring, maintaining, and improving a band of frequencies for the government.

The need for rental charges, as for shadow prices, should now be clear. At present, incumbents do not benefit from releasing spectrum to others even though necessary investment costs to reduce occupied bandwidth may be considerable. Yet their reticence to accommodate newcomers

is explicitly endorsed by the National Academy of Engineering in [29, p. 54]. The kind of auction scheme outlined below, however, appears in neither document, and differs substantially from the one described in [9] [35].

¹⁷ Only first option has thus far been explicitly considered in the literature. See Federal Communications Commission [11, Vol. 2, pp. 409-13].

imposes substantial costs on the latter, who, unable to share, borrow, or buy spectrum outright, are forced to develop systems higher up in the unused regions, or to substitute transportation, storage and personnel [21, pp. 342-43].

The incumbents' preferred status is grounded on their "continuing interest" in spectrum assignments due to prior use (regulatory doctrine to the contrary notwithstanding). It is based also on the difficulty of overriding the facts of past performance with the paper promises of future possibilities; and on incumbents' further political leverage if the service is an old one geared to safety, security, education, etc. However justifiable economically, the spectrum managers will rarely impose dislodgement costs. Could they raise rental charges, however, when the opportunity costs of occupied spectrum rose, this might make incumbents less reluctant to vacate, lend, or share spectrum to accommodate new services [21, pp. 346-49]. Long-term stockpiling would then be less necessary.

It is indeed hard to see why economists have so far dismissed the engineer's "spectrum usage measures"¹⁸ as irrelevant to the economic problem in allocational decisions. They have correctly pointed out the relation between spectrum value and the location of the bandwidth in spectral, geographic, and temporal terms. However, the true value of these measures lies in the base they provide for setting user charges.

Such charges could be imposed on the several dimensions of occupied bandwidth, with rebates allowed where users take steps to improve the efficiency of spectrum utilization. The more spectrum occupied in all major dimensions, the higher the rent. The less spectrum occupied, the lower the rent. Investment in new spectrum-economizing hardware will reduce the rent, but the spectrum user may prefer to keep his rent high by lax utilization to keep other costs low. He may thus trade spectrum for other factor inputs, or vice versa, just as he now strikes some balance among other factor inputs so as to maximize the net productivity of total dollar investment; or to equalize the productivity of money spent on all factors. The only difference is that the price paid for spectrum will not be determined in any full-fledged market.

C. Stratified (Zonal) Auctions. Under a stratified auction approach (within zones), market incentives and private valuations can be used without necessarily using them throughout the whole spectrum. Nor is there any reason why social val-

ues cannot be imposed over private values in specific areas. We would thus attempt to utilize market incentives within an overriding set of social constraints. But these overall constraints (the zoning rules) could themselves be modified later in light of the private valuations which emerge within predetermined zones, in turn reflective of social priorities.

One objection to any unstratified auction scheme is that the wealthiest bidders, in the major urban centers, are likely to preempt most of the rights at the expense, e.g., of rural areas where needs develop more slowly. It is feared also that local government, public safety, and educational or scientific users could never outbid private commercial entities in contests for spectrum. Here, too, auctions no less than freely transferable rights have been opposed on grounds of subverting vital regulatory priorities [19, pp. 90-92, 103-05].

But limits could certainly be placed upon the multiple ownership of rights by single entities, as well as on their geographic location. Frequencies could also be set aside for smaller, less affluent users to bid on. Local government, public safety, and educational users could be charged for their rights, too, but only required to bid against other like entities within common prescribed spectral zones.

In sum, stratified (zonal) bidding would require a specification of who is qualified to bid, where, against whom, and for what [23, pp. 29-33]. Bidder qualifications would be delimited in regard to geographic area, number of facilities already held, size of establishment, type of service, etc. But there is no reason why price incentives could not be injected into the current regulatory framework this way.

A second, closely related problem arises from the intensification of interference where disparate uses (say broadcast and point-to-point) are permitted within the same spectral region.¹⁹ Adjacent and harmonic channel interference and intermodulation problems may all be commensurately greater. For this reason, too, auctions might better be limited to like users within the same bands, with interband or interservice reallocations made administratively in the light of relative intraband auction values, but not uniquely determined by them. In land use, freedom to buy and sell does not mean freedom to buy "residential" land to build factories on, or vice versa. The market for land operates instead under the constraints of (albeit imperfect) zoning laws. Yet relative land val-

¹⁸ Representative measures of physical usage are described in JTAC [18, Sup. 8, pp. 85-93].

¹⁹ The possibilities of efficient sharing in other cases may have been unduly minimized in the past. See JTAC [18, Sup. 2, pp. 31-37].

ues within zones can and do point up the economic need for reallocation as between zones, such reallocation depending in the final analysis on legislative-administrative discretion wherein economic factors are tempered by sociopolitical considerations. The same could be done for the spectrum.

Purists contend, however, that bidding for rights across band or service lines is crucial for major improvements in frequency allocation. Surely the major conflicts today are less within than between current uses and bands. They lie in the heated contests between mobile radio and broadcasting, between military and nonmilitary or nongovernment users, or between common carriers and private users. Without interband reallocation the really big problems might never be touched.

Yet auctions across band lines are far more difficult to administer than intraband auctions. Even assuming that intermodulation and harmonic or adjacent channel interference could be handled by the proper federal designation of rights, it would be awkward at best to administer an auction between a thousand land mobile radio users and several TV broadcasters who want to lease the same (unoccupied) TV channel.

One way out is to authorize the numerous land mobile users to bid collectively for the TV channel, granting any necessary antitrust relief towards that end. One would presumably want: the broadcaster to hold out until the mobile radio users' offer price just equaled the full discount value of the spectrum to him; the mobile radio users to pay a sum equal to the extra costs imposed on them by exclusion. But the other more cogent approach, as already noted, would be to conduct intraband auctions within TV and land mobile; and then to use the relative auction values per megacycle bandwidth as *prima facie* evidence for the reallocation from one service to the other. Any subsequent cross-band reallocations would in that case take into account all relevant social considerations.

A third major contention is that benefit externalities and the public goods character of public, scientific, and educational radio services virtually preclude any viable auction scheme at all. Yet to say that public users neither could nor would pay for spectrum the sums necessary to retain what they now utilize may simply mean that: radio's value in implementing their organizational priorities is harder to measure accurately or to demonstrate persuasively to legislative committees for tax purposes than it is for private nonbroadcast users to do privately; or that even if radio's value could be measured with equal ease and accuracy

in both cases, its value to the public users is simply less.

The public users' fear of private competition undoubtedly reflects their difficulties in demonstrating for budgetary purposes the economic value of virtually any productive factor. Yet this hardly means that all proposed public uses of spectrum are necessarily more efficient in implementing public wants than private uses are in satisfying consumer preferences under competition. The fact that the public happen to own the spectrum and may choose to buy it from themselves and give it to the public users should be a matter of conscious public choice made in full knowledge of all the facts. Such public decisions would seem sounder if the community auctioned (or otherwise priced) the spectrum and gave the public users money to bid for it as for other factors.

Important merits notwithstanding, finally, auctions cannot do their job efficiently in the face of bidder collusion. Yet the very type of contest most likely to minimize such collusion by helping insure a large number of bidders is itself plagued by the most troublesome defects. Cross-band bidding in interband contests will doubtlessly be handicapped by the most intense problems of intermodulation and harmonic or adjacent channel interference. The reconstitution of rights may therefore be particularly complex and expensive. Yet many more bidders would appear in contests across band lines (than within them), and the danger of bidder collusion to hold down bid prices would also be smaller.

Intraband auctions under careful zonal constraints would be freer of the above-cited interference problems, but far more susceptible to bidder collusion. One has only to recall the small percentage of TV licenses granted through comparative proceedings and the relatively small number of candidates who compete even in those cases. Nor is it clear that dropping the minimal licensing standards would necessarily facilitate more applications or more bidders; nor that public officials would be willing to do so anyway.

By the same token, more explicit bidder qualifications and zonal constraints geared to safeguard regulatory standards would both operate to limit the number of effective bidders. The danger of bidder collusion would be commensurately greater. New safeguards will therefore be needed, perhaps the setting of minimum acceptable bid prices roughly related to shadow prices, since even crude shadow price values would be better than pulling the minimum values "out of thin air."

This latter point merely underscores our position throughout; viz., that some combination of

second-best options is the safest way to proceed during this formative period before all pros and cons have been carefully scrutinized.²⁰

REFERENCES

1. BAIN, CAVES and MARGOLIS, *Northern California's Water Industry* (1966).
2. COASE, MECKLING, and MINASIAN, "Problems of Radio Frequency Allocation" (unpublished, RAND Corp., May, 1963).
3. COASE, "The Federal Communications Commission," *J. of Law and Econ.*, Oct., 1959, pp. 1-40.
4. ———, "The Interdepartment Radio Advisory Committee," *J. of Law and Econ.*, Oct., 1962, pp. 17-47.
5. ———, "Problems of Social Cost," *J. of Law and Econ.*, Oct., 1960, pp. 1-44.
6. Communications Satellite Corporation, "Satellite Broadcasting: A Preliminary Report on Technical and Economic Aspects," Rev. B, July, 1967 (mimeo.).
7. DEMSETZ, "Toward a Theory of Property Rights," *A. E. R.*, May, 1967, pp. 347-59.
8. ———, "Exchange and Enforcement of Property Rights," *J. of Law and Econ.*, Oct., 1964, pp. 11-26.
9. *Electromagnetic Spectrum Management: Alternatives and Experiments*, G.E.-TEMPO, 68 TMP-64, prepared for the President's Task Force on Communications Policy and Office of Telecommunications Management. Transmitted Aug. 2, 1968, released Apr., 1969.
10. *Electromagnetic Spectrum Utilization—The Silent Crisis* (A Report by the Telecommunication Science Panel of the Commerce Technical Advisory Board, Oct., 1966).
11. Federal Communications Commission, *Report of the Advisory Committee for the Land Mobile Services* (1967), 2 vols.
12. *Final Report of the President's Task Force on Communications Policy*. Transmitted Dec. 7, 1968, released May 20, 1969.
13. ———, Rules and Regulations, para. 73.24.
14. HAVILAND, "Selected Studies of Space Broadcasting," *ITU Tel. J.*, Feb., 1965, pp. 77-85.
15. ———, "Space Telecasting for World Education," XVI IAF Congress, Athens, Sept. 13-18, 1965 (mimeo.).
16. Joint Technical Advisory Committee, *Radio Spectrum Conservation* (1952).
17. ———, *Radio Spectrum Utilization* (1964).
18. ———, *Spectrum Engineering—The Key to Progress* (1968).
19. JONES, "Use and Regulation of the Radio Spectrum: Report on a Conference," Washington Univ., *L. Q.*, Spring, 1968, pp. 71-115.
20. LEVIN, "Federal Control of Entry in the Broadcast Industry," *J. of Law and Econ.*, Oct., 1962, pp. 49-67.
21. ———, "New Technology and the Old Regulation in Radio Spectrum Management," *A.E.R.*, May, 1966, pp. 339-49.
22. ———, "The Radio Spectrum Resource," *J. of Law and Econ.*, Oct., 1968, pp. 433-501.
23. ———, "Regulatory Efficiency, Reform and the FCC," *Geo. L. J.*, Fall, 1961, pp. 1-45.
24. MECKLING, "Management of the Frequency Spectrum," Washington Univ., *L. Q.*, Spring, 1968, pp. 26-34.
25. MINASIAN, "Television Pricing and the Theory of Public Goods," *J. of Law and Econ.*, Oct., 1964, pp. 71-80.
26. MYERS, *Technical Aspects or Considerations of Frequency Assignment*, FCC Report No. F-6601 (1965).
27. NORTON, Annex of Letter from K. A. Norton to Arthur J. Goldberg, Aug. 14, 1965, analyzing letter from James D. O'Connell, DTM, to K. A. Norton, July 10, 1964 (unofficial multilith, ESSA/ITS, Boulder, Colo.).
28. PATTERSON, *A Survey of Techniques for Improving Utilization of the Radio Frequency Spectrum*, NBS Report No. 7630, Dec., 1962.
29. *Report on Selected Topics in Telecommunications*, Committee on Telecommunications, National Academy of Engineering (Final Report to Dept. of Housing & Urban Development, Nov., 1968; revised and submitted to Communications Task Force, Dec., 1968; released July, 1969).
30. ROSE, "Marketable Spectrum Rights," *Conference Record of the 1969 IEEE International Conference*, Boulder, Colo., June 9-11, 1969, Sess. No. 13, 7-12.
31. SAMUELSON, "Public Goods and Subscription Television," *J. of Law and Econ.*, Oct., 1964, pp. 81-83.
32. *Tentative High Frequency Broadcast Schedule*, IFRB (Geneva, Switzerland), Nov., 1967.
33. United Nations General Assembly, Outer Space Affairs Group, Working Papers on *Broadcasting from Satellites*, Feb. 13, 1969, and on *Broadcasting—Satellite Service: Technical Considerations*, Feb. 19, 1969.
34. U.S. Dept. of Commerce, *Ionospheric Radio Propagation* (1965).
35. DE VANY, ECKERT, MEYERS, O'HARA, and SCOTT, "A Property System for Market Allocation of the Electromagnetic Spectrum," *Stan. L.R.*, June, 1969, pp. 1499-1561.

²⁰ One final problem arises from the special difficulties posed by the reauctioning of existing licenses—in contrast with the initial auctioning of new grants. A careful demonstration of this point would take us far afield, but there is no question that competitive bidding for renewal rights is a far more complicated and potentially perverse matter than the letting of new grants for the first time [20, pp. 58-62]. So much so that public officials could well prefer to limit the auctions not only to applicants within the same bands, but to new grants as well.

DISCUSSION

ALFRED E. KAHN: One of the greatest obstacles to the efficient pricing of communication services seems to be a wide gap between the costs of service of the established common carriers and the costs of facilities embodying the most modern technology: microwave radio, satellite, and the newest coaxial cable. Levin's paper reminds us that this generalization applies clearly only to the comparison of private costs: until we take into account the opportunity cost of the spectrum that the first two of these new media use, we cannot be sure how great is their margin of superiority in terms of social costs.

But even with this correction, the rate of technological progress in long-distance communication has been so rapid it has apparently outrun the rate of depreciation permitted by regulatory commissions on existing facilities. In consequence, the embedded costs that the carriers are entitled to recoup from subscribers are apparently far higher than the costs of the newest facilities. This seems to be the main reason why, despite the presence of economies of scale, both Comsat and the far smaller carriers and private users mentioned by Leland Johnson have been able to undercut A.T.&T. and the other carriers on the bare line haul communication service.

It is of course at current, not historic costs that prices should be set in the interest of economic efficiency. It is only if rates are reduced all the way to current costs for those customers in a position to be served by facilities embodying the newest technology that that technology will be adopted at the optimum rate. In addition, the development of the technology itself is responsive to pressures of demand. Fulllest advantage can be taken of these dynamic economies of scale only if the pattern of rate reduction conforms to those potentials.

But this is impossible to do if, as Peck points out is typical, regulatory commissions permit the carriers to earn embedded costs. Nor is that policy unreasonable. It is not the fault of the common carriers if their past depreciation allowances have been unrealistically low. So their rates have been based on the average cost of a composite of new and old facilities and their captive customers have enjoyed the benefits of the newest technology only in proportion as it is actually adopted.

What's done is done. It seems unlikely that users can escape paying these sunk costs. As Peck points out, the proposed single international entity would probably have to buy out the cable investment, presumably at its unrealistically high book value; and rates would continue to embody a return on it. The hope is that we can avoid compounding these errors and distortions in the future, by getting the best possible investment decisions (and more realistic depreciation rates) from now on. And that is the major argument for a single entity in the international arena.

I do not know enough to offer a confident prediction that the chosen instrument approach will improve matters. Peck has ably reviewed the pros and

cons. I think he underestimates the Averch-Johnson danger if competition is entirely removed. It is not just that, as he puts it, "a single entity is not inherently a solution to the problem of rate base expansion"; the danger is accentuated if competition is eliminated. As he correctly observes, the most important mitigating factor is regulatory lag: the longer the delay in adjusting rates to altered costs, the greater the incentive of regulated companies to use the lowest-cost technology even when it is less capital intensive. But note that the FCC's continuing surveillance policy, justly hailed as a method of making regulation more efficient, has the unfortunate effect of shortening the regulatory lag—hence of strengthening the Averch-Johnson tendency.

On the other hand, I find persuasive some of the counterconsiderations that Peck adduces, explicitly or by implication. So long as Comsat is confined essentially to the role of a carriers' carrier, the only real competition possible between it and the other carriers is for certification. Faced with competing applications, the FCC is under strong temptation simply to divide the market between the applicants. Note, for example, the fair shares policy embodied in its 1968 decision authorizing the terrestrial carriers to lay down another transatlantic cable—the TAT-5¹—when it attached the condition that the parties use satellite circuits in sufficient numbers to fill the unused satellite and cable facilities "at the same proportionate rate." And note its refusal in that case seriously to decide which method of meeting the incremental demand would be more economical: "We do not believe that any useful purpose would be served by going over relative costs. . . . We do not feel it necessary to make definitive findings on the relative merits of TAT-5 and present . . . satellites. . . . There are difficulties in making comparisons between cable and satellite costs."

And, as Peck suggests, it could well be that constant scrutiny by the large users of the international circuits, led by A.T.&T., would be a far more effective means than the limited kinds of competition that prevail today of assuring that the chosen instrument did in fact employ the lowest-cost combinations of facilities.

But it is very hard for an economist to give up even the present limited possibilities of competition. The opportunities for dynamic loss in any situation that gives a single company discretionary control over what directions technological change is to take, subject only to regulatory scrutiny, could well exceed the very great efficiency gains that a single, multimodal carrier could achieve. Perhaps the better alternative is to permit more direct competition between Comsat and the terrestrial carriers, by permitting more of the

¹ In the matter of A.T.&T. *et al.*, *Applications for authorization to participate in the construction and operation of an integrated submarine cable and radio system*, etc., Memorandum Opinion, Order and Authorization, 13 FCC 2d, 235 (1968).

users who can deal directly with it to do so, just as the FCC has authorized direct competition between private and independent common carrier microwave radio systems and A.T.&T. This kind of competition can only be selective: only the large users can take advantage of it. And this means that the benefits of the new technology would accrue in full measure only to those users. But the competition would be genuine and would come closer than any other device to getting rates—some of them right away, all of them in time—down to marginal costs, which is where they belong.

The essential characteristic of the public utility is the limited possibilities of competition. The growing rivalry that Johnson describes is highly selective. If, as he correctly points out, the existing carriers are not permitted to meet it by reducing rates discriminatorily down toward their long-run incremental costs, serious inefficiencies may be introduced, depriving all subscribers of the benefits of economies of scale. But the result is discrimination—and unless the competition is carefully controlled, it could result in higher rates for the captive subscribers than they would otherwise pay. It may well be also that the new domestic entrants are really being subsidized by those captive customers in another way: to the extent their cheaper and in some ways lower-quality and less-reliable service is attractive to customers only because A.T.&T. provides them backup—if the proffered service fails, the customer can always pick up his telephone—the regular subscribers may in fact be subsidizing them by bearing the total cost of those backup facilities.

But freer competition need not be wholly discriminatory. On the contrary, it may force the existing common carriers to introduce a less discriminatory pricing system than they now employ. The principle of averaging that underlies their rate making extends to the costs not just of new and old technology but also of the various geographic parts of their operations. A.T.&T. prices its services uniformly on the basis of nationwide or system-wide average costs. Since the tendency to decreasing costs is most marked when given facilities are used with greater intensity, it appears that the company's cost of service is far less on the intensively utilized, heavily traveled routes, and corresponding far above average on the thin routes. Yet the rates are the same per mile on both. This apparently has been one factor facilitating the entry of competitors on the former routes. The FCC has never clearly confronted this problem. If it wants A.T.&T. to price uniformly, then it will have to preserve the system of internal subsidization that this involves by denying competitors the right to enter the lower-cost-higher-density markets, as the dissenting commissioners in the MCI case contended it should have done. If instead it wants users to enjoy the benefits of competition, it must also envisage the breakdown of uniform, nationwide pricing and permit A.T.&T. as well to move to a system of nondiscriminatory, route-by-route pricing, if important economies of scale are not to be lost.

So competition of sorts is feasible in communications, both domestically and internationally. But, par-

adoxically, its vastly increased possibilities of the last decade also increase the importance of regulatory scrutiny—to see to it that discriminatory price-making does not in fact end up imposing burdens on the captive subscribers who do not enjoy its benefits; and that competitors do not get a free ride on the facilities of the switched national network or impair the quality of its performance. And finally, as Levin demonstrates, competition can be grossly distorted if the various rival suppliers do not have to pay the opportunity costs of the spectrum that they use.

ROGER G. NOLL: The Chairman of this session has requested that I focus my remarks upon Professor Levin's paper on allocating the electromagnetic spectrum.

The most telling argument against proposals to establish a market in spectrum rights has been that the nature of spectrum interference makes a free spectrum market unworkable. Levin accepts this argument, but he then goes on to show that a limited market system could be beneficially established in many regions of the spectrum even if the interference problem is as difficult as opponents of markets have contended.

In my view Levin overestimates the importance of the interference problem and therefore underestimates the extent to which it is technically feasible to develop a spectrum market. Levin argues that a reasonably unfettered market in spectrum rights would not function efficiently because of two aspects of the spectrum interference problem: the unpredictability of spectrum interference, and the existence of an interaction effect between spectrum emissions, causing interference to a third user when neither interfering emission would, by itself, cause interference. Spectrum interference occurs most frequently between users of adjacent bands, harmonic frequencies, or the same band but in different geographical areas. These types of interference are unpredictable in the sense that the degree of interference depends partly upon random, uncontrollable, natural phenomena. But the probability of these phenomena occurring is reasonably well known, and the interference problems they create can be controlled by altering signal "inputs," such as signal strength and directionalization, antenna height, or the specifications of signal-generating equipment. This knowledge is the basis for existing spectrum allocations by the FCC.

An unregulated market in spectrum rights as currently defined (but made infeasible) would not work, since technological advances would soon create a greater interference capacity within a given specification of input limitations. But an unregulated market could work if spectrum rights were defined geographically and in probabilistic terms; i.e., a user could radiate on a given band so long as his received signal strength did not exceed some maximal level anywhere outside the perimeter of his geographical area more than some specified fraction of the time. Among the probable consequences of "output" regulation of the spectrum would be the development of a new industry of spectrum monitoring. Firms would place devices on the boundaries of spectrum-use areas which would record the received signal strength from

all users whose spectrum rights were confined to one of the adjacent geographical areas. Users would then use information from the monitoring device to vary signal-generating components to keep received signal strength within the specified limits. Such a system would result in less interference than the present system, for users would have added incentive (the possibility of legal action and/or fines) to alter emitted signal strength in response to changes in natural conditions affecting signal reception.

Levin's argument about the difficulties of dealing with interaction interference has more technical validity, but fortunately this phenomenon is very rare, and probably not important enough to be a governing factor in the choice of spectrum allocation institutions. In the rare instances when such interference does occur, the cause is usually in the proximity of antennas. In this case it is rather easily solved by re-spacing antennas or by constructing shields and traps around existing antennas. Because of their rarity and their rather inexpensive solution, cases involving this type of spectrum interference would be unlikely to reach the courts, but if they did, then responsibility for correcting the situation would probably be placed upon the most recent entrant among those mutually interfering. As long as the courts allowed the responsible party to effect the least expensive solution that did not violate anyone's spectrum rights (e.g., the responsible party could construct a shield for another antenna as well as move his own), no serious inefficiencies or inequities would arise.

The problem that most clearly could not be handled by a market and access to the courts is that of spurious emissions. This type of interference generally occurs over too narrow a geographical area for a single instance to create damages sufficient to justify lawsuits, yet the cumulated damage of all spurious emissions could be high in an unregulated environment. The easiest solution is to set national standards limiting the radiation generated by electrical machinery, but this creates the possibility of significant inefficiencies. There is no reason to require the same standards for electrical equipment located in unpopulated areas as are required in big cities. Yet regulation of spurious emissions probably can be efficient only if applied uniformly at the production stage. Furthermore, the probability of a spurious emission creating interference depends upon the technical composition of the receiving device and the geographical separation of sender and receiver as well as upon the strength of the emission. (Receiver quality occasionally is a factor in interaction interference, also.) The optimal solution to many spurious emission problems is to improve the quality of the receiving device: a ham operator who interferes with his neighbor's television reception will usually place a trap on the TV rather than improve his own equipment. Because the spurious emissions problem is so knotty, a market system probably cannot protect spectrum users from this type of interference, and direct regulation is therefore required.

Another major difficulty of a market system relates to radio and television broadcasting. The profitability

of purchasing spectrum outside the present broadcast band is dependent upon the ability of the public to receive the broadcast. The development of UHF television was retarded by the reluctance of manufacturers to build television sets that could receive the newly opened channels. After the FCC required that manufacturers include UHF reception capacity in all sets, the development of UHF was governed by the growth of UHF receivers in use. Without regulation, manufacturers will not expand the reception capabilities of radio and TV sets until broadcasts on the new band generate adequate consumer demand; yet broadcasters will not use the new bands unless receivers are equipped to pick up their broadcasts. If a spectrum market is to work in allocating the spectrum among different types of users, the FCC must retain the authority to define the band reception capacity of receiving devices. This conclusion is not without normative overtones: if an individual knows that he will never want to watch UHF television, why should he be forced to pay more for a television set capable of picking up worthless (to him) broadcasts?

The preceding discussion suggests that a workable market for the spectrum is technically feasible; however, normative considerations may encourage us to shy away from maximal dependence upon the market. Aside from the argument Levin puts forth with regard to spectrum use by public agencies, the development of an unregulated spectrum market could have an important impact upon the quality of broadcasting. Currently broadcasters have some incentive, though admittedly little, to produce "public service" (i. e., less profitable) broadcasts to satisfy the "public interest" clause of the Communications Act as interpreted by the FCC. An unregulated market for broadcast spectrum would remove this incentive and would also insure that only the more profitable bidders for spectrum obtained radiation rights. Under these conditions, one would expect even less variety, quality, and relevance in broadcasting than currently exists.

Probably the only hope for improving television quality is to increase competition in the industry. As long as the number of television channels (and, particularly, networks) is small, programming will ignore the tastes of groups too small or too poor to exercise significant effective demand for an advertiser's product. Making broadcasting a private good, either by garbling signals or by switching television to wire, will add effective demand for program content to the factors influencing broadcast decisions, creating incentives to program for smaller but affluent groups of viewers. The "wired city" proposal has the additional advantage of freeing substantial spectrum: a television station uses several hundred times as much spectrum as a typical land mobile user. Removing television from the spectrum in large cities would, for a while, make the spectrum a free good again in almost all of the United States. This fact is not incidental to the prospects of achieving some changes (perhaps a market) in spectrum allocation. The prospects of establishing a spectrum market are improved if the market price will initially be very low, thereby assur-

ing that a new allocation procedure will not produce an immediate reshuffling of radiation rights and a loss of investment by some users. Unfortunately, removing all television from the spectrum has two serious drawbacks. First, switching totally to pay TV will have a highly regressive income redistribution effect: lower income groups watch television more than others, but spend less on advertised consumer goods. Second, one of television's most useful functions—allowing mass visual communication on issues of great national importance—would be eroded.

The preceding analysis suggests that normative considerations and not interference problems are the binding constraints on the extent to which a market system can be used for spectrum allocation. At the same time, the differences between Levin's conclusions and mine should not be overemphasized. We agree that much greater reliance should be placed upon market devices. Together the two papers indicate the bounds within which feasible market options lie.

WILLIAM H. MECKLING: The papers prepared for this session are a welcome diversion from the usual fare at these meetings. All truly address important current policy issues. Peck and Johnson deal with market structure in the communications industry, while Levin addresses the management of the frequency spectrum.

Peck reviews the rationale for establishing a single entity to supply overseas transmission services. He is concerned that multiple ownership will lead to parallel expansion of cable and satellite facilities, whereas cost considerations dictate that any expansion should emphasize satellites. Unfortunately, the abbreviated nature of his cost comparisons and his descriptions of how the estimates were derived leave some doubt about their validity. More important, in discussing the advantages and disadvantages of a single entity, he neglects the long-run entry problem which characterizes a regulated single entity. Moreover, he offers no analysis or evidence to support the view that a single entity would, as opposed to could, be regulated so as to provide more effective overseas service even in the short run. The will and power of the Federal Communications Commission to regulate domestic communications is open to question. On the international scene where it confronts an array of government-owned communications enterprises, the Commission's ability and willingness to regulate are virtually nonexistent. A single entity in overseas communication means an unregulated monopoly. This is, of course, precisely the arrangement which prevailed pre-Comsat. Before we discard the increment in competition, however small, which Comsat has introduced, it would be comforting to have more assurance that the prices charged for overseas service by the new single entity will be more sensitive than those which prevailed for two decades after World War II.

Mr. Levin is anxious to improve the management of frequency spectrum, but rejects what he calls a "full-fledged" spectrum market as the solution. He lists three characteristics of frequency utilization "that impede the efficient definition, enforcement and recon-

stitution of rights in spectrum." These are: (1) the variability in signal outputs due to natural phenomena beyond the users control; (2) the general unconfinedness of radiation irrespective of spectral region or transmission mode; (3) the intermodulation effects. These three factors are alleged to raise questions as to whether rights can be bought, sold, or reconstituted efficiently in a bona fide market. Levin combines this analysis with political considerations and on the basis thereof recommends a "middle option" as the best strategy for frequency management, where middle options range from shadow pricing to the auction of federally designated nonresalable radiation rights.

Anyone who has worked on the frequency management problem must be impressed that doing anything constructive is at bottom a political problem. Levin's attempt to suggest a staged approach to implementation of a market is in that respect to be commended. Nonetheless, his analysis is distressing, and his recommendations neglect what seem to me to be the most important political realities.

The distress stems from having an economist needlessly lend aid and comfort to the enemy; namely, the engineers and physical scientists, who for ten years have been raising the same specters Levin raises as reasons (among others) a market for frequencies won't work.

They are not valid objections to a market for frequencies. They confuse the question of whether some "ideal" private rights can be defined and exchanged with whether any rights can be defined and a market for them created. It should be clear that radiation rights need not be defined in terms of emission or output; they can be defined in terms of inputs just as they are now. From that standpoint the fact that cries out for attention is that we already have a market for rights. The right to radiate energy is bought and sold all the time in the broadcast bands. In view of that fact, it is a little hard to understand what people mean when they say that rights cannot be defined, so as to effect a market for them.

Second, even if rights are defined in terms of emission or energy output (which I believe to be preferable), it is not necessary that such rights in some sense be perfect. Rights are not perfect in other resources; e.g., the view problem and land. Is that a valid reason for abandoning the market for land? We are confronted here with the problem so prevalent in economics today, that it is not sufficient to show that a market is better than any alternative; one must demonstrate that it is perfect. All three of the problems raised by Levin exist whether there is a market for frequencies or not. What do we do about these three problems now? The answer is virtually nothing, and insofar as anything is done, e.g., enforcing manufacturing standards for noncommunications devices which emit energy, those same things could be done under a market system. With private ownership of rights at least, there would be an incentive to develop means of detecting and curtailing spurious emission, and for forcing those who radiate to adjust their out-

puts to variations in natural phenomena. We might even find out more about the natural phenomena. In short, there may be reasons for a stepwise approach to creating a better market for frequencies, but if so, that has nothing to do with the feasibility of a market system.

On the political side, it seems to me that Levin's suggestions ignore some of the most potent political facts of life. For example: (1) The present system of allocation and assignment has generously endowed powerful interest groups with emission rights; e.g., the airlines, hams, radio astronomers, the military services, broadcasters, the American Telephone and Telegraph Company, and on and on. Those groups are not going to take lightly the threat that Congress or the FCC will impose user charges for the exercise of rights which they have had at no cost up to now. (2) Politicians, generally, and Congress in particular have a special interest in the broadcast business. Many politicians and congressmen are part owners in television and radio stations. More importantly, even "nonpolitician" owners of broadcast facilities have potent political power. Frequently, local newspaper publishers also own a TV or radio station. No congressman who is seriously concerned about staying in office is likely to vote for a measure that requires these organizations to begin paying for a resource they have heretofore been authorized to use without charge.

Given these powerful interests, it would seem more practical to think in terms of bribes we can offer than charges we can levy if we hope to make any progress. If we could combine bribes with a few more speeches by Mr. Agnew, we might persuade the industry of the merits of converting frequency spectrum to private ownership.

WILLIAM K. JONES: All three papers raise a common question: what should be the role of regulation, as opposed to competitive market forces, in shaping national telecommunications policy? I would like to make a few observations about the nature of regulation.

First, most regulators take their jobs seriously, and it is extremely difficult to persuade any of them that the functions they are performing are contrary to the best interests of society. Thus, the frequency manager (the principal subject of Levin's paper) and the common carrier regulator (the principal subject of the Peck and Johnson papers) consider themselves to be endowed with something akin to a holy mission. They will not easily relinquish the authority to allocate frequencies to their "proper" uses; to select the "best" users of those frequencies from among competing applicants; to prevent the "unwarranted" entry of newcomers; and to set aside "unfair" reductions in rates. To do so would be a form of regulatory heresy.

Second, most regulation creates a host of special equities (or vested interests) which make the introduction of market pressures especially difficult. The spectrum user, of course, is the most flagrant example of somebody who "is getting something for nothing" and who will seek to preserve the regulatory scheme that provides such a bonanza. The Levin paper devel-

ops this problem in detail. Even those who are relatively disadvantaged—such as land mobile users and communications satellite proponents—shy away from the idea that the appropriate method for obtaining greater access to the spectrum is to pay the government for the privilege (in amounts exceeding those that alternative users would be prepared to pay).

The Johnson and Peck papers cite some more run-of-the-mill examples: the franchised firm protected against competition by certification requirements; shareholders with a stake in obsolescent investment; high-cost customers who obtain services at reduced rates because the regulated company is in a position to charge other customers rates considerably in excess of cost. All will vigorously resist entry and innovation which threaten their interests.

Third, Congress consistently gives regulatory agencies lofty missions to fulfill but is extremely niggardly in allocating the resources needed to accomplish those missions. There are some talented individuals among the regulators, but they are few and far between. In my opinion, the problem of governmental personnel will get worse before it gets better. Does Levin really envisage a Federal Communications Commission or an Office of Telecommunications Management capable of computing and utilizing "shadow prices" with any degree of sophistication? Does Johnson believe that the FCC will surpass the ICC in its capacity to properly compute, and wisely use, such concepts as long-run incremental costs and fully allocated costs—these have been the grist of the ICC's mill for over thirty years in its efforts to control rail-motor competition? (I gather from Johnson's conclusion that he does not.) Does Peck believe that the FCC will be in a better position to cope with a "single international entity" than it has been in respect of A.T. & T. on the domestic front? A.T. & T. is much larger than any international entity would be, but size is not dispositive; it is the availability of alternatives that counts. And for the international entity there would be none. Moreover, it is unclear how any government agency—including the FCC—can control the rates for the "space segment," which are fixed by Intelsat and not the FCC, unless there exists some cable capacity which can be employed as leverage in preventing excessive rates.

Fourth, the combination of the preceding factors—the concern of the regulators about their prerogatives, the vested interests or equities created by regulation, and the inadequacy of regulatory resources—all combine to create enormous pressure for the preservation of the *status quo*.

In the usual instance, there is the problem of inordinate delay in obtaining necessary approvals from regulatory authorities. Where regulatory prerogatives are threatened, or vested interests among the regulated are disturbed, there may be a flurry of activity—as, for example, in the case of CATV. Otherwise, proceedings tend to drag on and on. The domestic satellite proceeding is perhaps the most notorious example. But even the MCI case, to which Johnson points, is still pending on petition for reconsideration;

and the scope of the "foreign attachments" opinion is still being explored in the context of the complex tariff revisions of the telephone companies which are being hotly debated.

In decisions on the merits, the interests of existing companies receive enormous weight. This is true in the area of spectrum allocation, as noted by both Johnson and Levin. And it is evident in the "traffic sharing" decisions involving cable and satellites (Puerto Rico and TAT-5) to which Peck refers. There are exceptions, to be sure, and some of them are quite important: the opening of the spectrum to private microwave; the loosening of the rules relating to foreign attachments. But these are exceptions and they required enormous struggles—with the backing of powerful industrial interests (the private microwave users and the manufacturers of electronic equipment).

These problems, to me, suggest a greater reliance on market forces.

The issue confronted by the Levin paper is the most intractable. All four of the factors I have mentioned are lined up in opposition to any movement toward greater emphasis on market forces. More in desperation than in earnest, I propose that frequencies be leased at public auction for uses predetermined by the FCC; that Congress direct the FCC to maximize revenues from the leasing of the radio spectrum; that 99 percent of the proceeds so raised be employed in a "package" of educational, welfare, and urban programs that might hopefully attract allies; and that the remaining 1 percent be paid into a pension fund for retired congressmen, Federal Communications Commissioners, and FCC staff.

On the Peck paper, there is little that I can add to the proposal I made to the staff of the Presidential Task Force: that the problem of cable construction be resolved by a publicly announced policy to the effect that A.T. & T. should be permitted to maintain its present capacity in cables but should not be per-

mitted to increase its capacity unless it divests its cable facilities to a separate cable company. If a separate cable company can survive in competition with satellites, A.T. & T. should have no greater objection to divestiture here than in the case of the single international entity (which would require similar divestiture). If a separate cable company is not viable, there is no substantial case for expanding existing cable facilities. Comsat, which affords the greatest promise for technological progress and economical international communications, should be cut loose from the conventional carriers; it should be permitted to deal with ultimate users (i.e., the "authorized user" decision should be reversed); and it should be given access to such users, where local telephone monopolies intervene, through compulsory interconnection.

For me, the Johnson paper presents the most intriguing discussion, since it presents in a microcosm the basic question of whether it makes any sense, as a general matter, to have a regulatory scheme characterized by certificates of public convenience and necessity and minimum rate regulation. I have my doubts. Special problems may be presented by spectrum utilization and vertical integration, but generally it seems that, if an incumbent company can do a job more efficiently than a newcomer, it does not need regulatory assistance in excluding the newcomer. And if an incumbent company cannot do the job more efficiently than the newcomer, it does not deserve, and should not be given, regulatory protection. I would leave any problems of predatory pricing to the anti-trust laws (admitting that they are not wholly perfect in this respect), and would leave service to rural areas to such agencies as the REA and other direct support programs. But I believe that the Johnson paper—and the other papers as well—shed valuable light on the important problem of government licensing of new entry in the regulated industries generally and in the telecommunications field in particular.

ECONOMICS OF IMPERIALISM

MODERN IMPERIALISM: THE VIEW FROM THE METROPOLIS

By RICHARD D. WOLFF

City College of the City University of New York

This essay aims to assess the role of imperialism in the economic structure of modern, advanced capitalist nations, particularly the United States. For much of the discussion, my vantage point will be that of the largest oligopolistic corporations which most economists view as comprising the typical locus of decision making—at least on the production side of such economies. Given the shares of total employment, assets, sales, and investible funds usually accounted for by, say, the 500 largest of such corporations, an assessment of the significance of imperialism for them may serve as a first approximation to its significance for the economy as a whole.

¹ Between 1870 and 1914 the advanced capitalist economies all experienced intense public discussions on the question of imperialism which then meant specifically the acquisition of colonial possessions. Since that time, imperialism has taken on a broader significance. The term has been used generally to describe the network of means of control exercised by one economy (enterprises and government) over another. Colonialism is then only one such means, decreasingly in evidence as time passes. Multinational corporations and foreign aid, for example, are among some other major means of control.

Control of part or all of a foreign economy is an increasingly important objective of oligopolistic corporations and their governments in advanced capitalist economies. The corporate striving to maximize profits (or sales or growth) implies and is paralleled by a striving to secure maximum control over parts of or entire foreign economies. Of course, concrete and changing historical circumstances may at any moment render the achievement of maximum or even any profits impossible. So it is with the achievement of the imperialist objective of control. Conditions within or among oligopolistic corporations and their governments on the one hand and the less developed nations on the other may make the degree of control attained vary greatly from case to case and from time to time. However, what persists in today's international economy is the ceaseless striving of the oligopolistic corporations and their governments to establish and expand control.

Control is sought by an enterprise and/or its

government to assure the secure provision (i.e., not subject to the uncertainties and "imperfections" of the market) of the following: (1) essential imported raw materials and foods; (2) markets for its manufactured exports; and (3) spheres for the investment of its capital.¹ Thus modern capitalist imperialism comprises a complex of private corporate policies, supplemented by induced governmental support, seeking to develop secure sources of raw materials and food, secure markets for manufactures, and secure outlets for both portfolio and direct capital investment.

I

World trade in the era of free trade in the mid-nineteenth century and subsequently in the era of imperialism developed a complex international flow of raw materials.² Great Britain carried the world division of labor to its furthest degree by her total dependence on imports of cotton and heavy dependence on imported grain. However, to varying degrees enterprises in all the capitalist countries sought, found, and developed foreign sources of raw materials and foods. In the last 150 years the availability of raw materials and foods coming from abroad has had cumulative and far-reaching consequences for the metropolitan economies.

Basically the availability of any specific raw material has in the past and continues in the present to determine key aspects of the development of capitalist economies. First, firms using such materials as essential inputs will undertake investments and thereby stimulate secondary investments according to their experiences with—as well as their expectations of—the continued availability of the materials at some specific range of prices. Second, other firms that establish transport facilities, warehousing, and distributional outlets necessary to utilize the raw materials will

¹ For a comparable discussion of the causes for and benefits to corporations controlling or evading domestic market mechanisms, see J. K. Galbraith, *The New Industrial State* (Houghton Mifflin, 1967).

² The argument regarding foods is similar to that for raw materials. For reasons of space, I will not develop it here.

make expenditures which themselves then stimulate further secondary investments. Over time, the profitability and hence the pattern of all such investments will depend significantly on the terms of the materials' availability.

Third, perhaps the most fundamental influence of the availability of imported raw materials lies in the historical process of technological change. Technological innovation proceeds from some presumed degree of availability of raw materials as well as the technical problems requiring solution. In this sense, the availability of imported raw materials will condition the direction of technical change.³ Since every technical change is more or less predicated on related changes achieved at a prior time, expectations about imported raw materials become embedded in the importer's technology. The availability of such supplies at any moment of time thus enters as a determinant into the heart of the productive processes of the capitalist economy.

For one illustrative example, consider the modern jet engine. The series of technical innovations that produced this item have made its production absolutely dependent on columbium, chromium, and cobalt—all commodities whose consumption in the United States is totally dependent on imports. Or consider that in 1966 net imports of iron ore equaled 43 percent of domestic United States production.⁴

In summary, capitalist enterprises develop a cumulative adaptation to the availability of raw materials including those imported. Such adaptation implies some dependence on the part of private capital. Enterprises function around specific expectations of prices and quantities of imported raw materials. In this connection, I wish to stress that while changes in tastes or technology may lessen any industry's dependence on some imported raw materials, as long as any quantity still needs to be imported for any production, the state of dependence continues.⁵

As cumulative adaptation implies dependence, so dependence implies anxieties on the part of cap-

italist enterprises. Their managements confront three major possible interferences with the desired quantities and/or prices of imported raw materials. Any particular foreign sources may be naturally depleted. Or new sources may fall under the control of some other management group seeking exclusive or preferred access to the raw materials. If the new source yields significant price or quantity advantages, the market shares (and hence the profits and growth) of enterprises with or without such access to these imports may be sharply altered. Third, control over existing or new sources may fall to local political forces bent on wage policies, etc., that would change the quantities and prices of their exports to the detriment of their buyers. In conventional terms, the uncertainties of commodity and capital markets and various actual or potential market imperfections do in fact threaten the stability and status of enterprises dependent on imported raw materials.

Responding to any or all of these anxieties, corporate managements generate multifaceted strategies. First, where possible they seek to locate, secure, and develop alternative sources of needed raw materials within different political sovereignties to spread their risk. The tactic here will often be vertical integration. Another tactic may see oligopolists of one nation or international grouping acting in concert to stimulate alternate sources and divide the known international sources amongst themselves through a more or less coordinated monopsonist approach. Under all circumstances, singly or in groups, private enterprises seek to utilize the international power of their governments to assist their efforts to overcome existing or possible threats to their lines of dependence. Their objective is to ensure that political authorities in source countries pursue policies which at least block any interference with the desired terms of raw materials exports. Further, as much of the world as possible should be open for the private development of new, perhaps as yet unknown, raw material sources. In pursuance of these objectives, the political complexities of contending groups in all countries become matters of direct and continuous concern.⁶

Searching at least to secure and at best to advance their market positions, capitalist firms and

³ At the end of the nineteenth century, Great Britain had ample access to natural dyes. Germany did not. The consequent economic history of the synthetic dyestuffs and the derivative chemical industry took very different directions in the two nations with important differential consequences. See R. Sasuly, *I. G. Farben* (Boni and Gaer, 1947).

⁴ See H. Magdoff, *The Age of Imperialism* (Monthly Review Press, 1969), pp. 48–52, and works cited therein.

⁵ See data cited in footnote 4 for evidence on the continued U.S. dependence on imported raw materials.

⁶ A. A. Berle, Jr., carries this point further when he writes of corporations that "have continuous and careful reports made to them on the attitudes and aptitudes of the American diplomatic officials, rating them according to their probable usefulness in advancing or protecting the company's interests." *The Twentieth Century Capitalist Revolution* (Harcourt, Brace, 1954), p. 132.

governments have moved to establish control over foreign economies directly, through vertical integration and colonialism in some cases, and by less overt pressures on and manipulations of markets and formally independent governments in other cases.⁷ Amidst changing conditions and with varying degrees of success, they have persisted in their striving to maximize effective control over needed raw material sources. In this connection it is worth noting that the most concerted effort of a mature capitalism—that of Nazi Germany—to achieve security not by controlling and protecting lines of dependence but by breaking them, was on the whole unsuccessful.⁸

II

The levels of sales achieved by corporations have always been key variables in the objective functions of managements. Whether sales were ends in themselves or means to higher profits, business has generally sought to secure the widest possible markets for its outputs. The history of capitalism and particularly its dispersion across the globe are heavily bound up with foreign sales, especially of manufactured goods. Moreover, the oligopolization of industry with its concomitant limitation of the domestic market has intensified managements' interests in foreign sales outlets.

Whether or not we assume that capitalist economies only export what was first produced for domestic demand, a foreign market that may once have been marginal or accidental will often provide sufficient profit to induce the firms concerned to make efforts to secure that market. Such efforts in themselves will likely include elements of what is here called imperialism. In any case, when such efforts prove successful and consequently an initially marginal foreign market becomes a "normally expected" market, the affected corporate managements will make investments dependent for their profitability upon the continued fulfillment of normal expectations of foreign sales. The protection of these lines of dependence on foreign sales acts as a major spur to imperialism in advanced capitalist economies.

Nor is there concern only with the protection

of existing foreign sales at any moment. National or international agreements among oligopolists over market shares—presumed to reflect underlying economic strengths—may be upset and/or revised if member corporations have differing degrees of access to rapidly expanding markets (or to raw materials sources as discussed above). Thus corporate managers confront both the carrot and the stick of foreign sales. If an actual or potential competitor has preferred access to an expanding market, his relative economic strength and hence his share of a market may rise. Similarly, the development of special advantages in a market that is now or may eventually be relatively buoyant must attract the attention and expenditures of ambitious managements; e.g., the rush of U.S. corporations to Western Europe. To the objective of securing existing foreign markets we add that of developing the most advantageous possible access to potentially important markets. As in the case of raw material imports, firms move with their governments to exert control over foreign markets.

We have used sales in this discussion rather than exports to encompass the increasing moves of particularly U.S. corporations to supply markets from foreign producing subsidiaries. Access to relatively cheaper labor or other inputs will frequently motivate such movement. Often the movement follows from a desire to keep export markets when they are closed or threatened by disadvantageous tariff or similarly restrictive regulations. Indeed, corporate strategy may establish subsidiaries, which make low profits or even losses for a while, when the larger objective is the maintenance or development of a firm position in a particular foreign market. In general, the establishment of such subsidiaries is another stage in an evolving corporate strategy which, by securing lines of dependence, simultaneously deepens them.⁹ Policies aimed to protect actual or potential exports will only be intensified with the addition of foreign producing subsidiaries to the overall picture.

I can summarize that picture with rough data for 1968 which indicate that beside some \$35 billion in U.S. exports, U.S. investments abroad—mainly direct investment in subsidiaries—ac-

⁷ Cf. R. Robinson, J. Gallagher, and A. Denny, *Africa and the Victorians* (St. Martin's, 1961). This work argues that British colonialism in Africa was in large part designed to safeguard her economic lines of dependence on India.

⁸ Germany after 1933 made massive investments in the direction of a much-desired autarky. The failure to achieve autarky contributed to fascist Germany's imperialist policies in the later 1930's. See B. S. Klein, *Germany's Economic Preparations for War* (Harvard, 1959).

⁹ For concrete examples in the recent history of a corporation, the Ford Motor Company, and a country, Brazil, see M. Wilkins and F. E. Hill, *American Business Abroad* (Wayne State Univ. Press, 1964), pp. 406-32; and C. McMillan, Jr., R. F. Gonzales and L. G. Erickson, *International Enterprise in a Developing Economy* (Michigan State Univ. Business Studies, 1964), pp. 28-33.

counted for \$200 billion in further sales.¹⁰ The importance of foreign subsidiaries for the U.S. economy generally, as well as for individual corporations, is also clear. International investment, it appears, has grown twice as fast as world GNP in the last twenty years. The role of U.S. corporations in that investment has risen rapidly over the same period.¹¹ Active U.S. corporate participation in this very dynamic sector of the world economy continues to offset certain internal tendencies toward slower growth rates and stagnation.

III

Historically, differential rates of interest presumably motivated foreign portfolio investments. However, foreign direct investment has traditionally been a major component of corporate strategy aimed at securing the raw materials, food, and markets described above; that is, foreign direct investment has often been a necessary means to protect lines of dependence. However, the very success of businessmen in securing their lines of dependence historically accelerated the oligopolization of industry which in turn generated another, related but distinct, pressure for foreign private investment. This is the so-called "surplus capital" basis for capitalist imperialism.¹² The argument on this point emerges from the simple analytics of oligopoly.¹³ At any given time, oligopolistic corporations, particularly now in the United States, confront generally limited and shrinking investment opportunities within their own lines of business. Retained earnings begin to be shifted increasingly away from existing lines of production into other, as yet unoligopolized, lines, into new products, and into foreign investments. The adjectives conglomerate and multinational are increasingly applied to such corporations. In simplest terms, the reason for direct investments abroad—as it appears to the investors—is the higher expected marginal rates of return abroad.¹⁴

¹⁰ Estimates offered in Judd Polk, "The Internationalization of Production" (U. S. Council of the International Chamber of Commerce, May 7, 1969, mimeo.).

¹¹ *Ibid.*

¹² Cf. V. I. Lenin, *Imperialism, the Highest Stage of Capitalism* (Moscow: Foreign Languages Publishing House, n.d., Chap. IV). Lenin did not limit his discussion of causes of imperialism to the surplus capital argument. See also P. A. Baran and P. M. Sweezy, *Monopoly Capital* (Monthly Review Press, 1966), pp. 178–217; and T. Kemp, *Theories of Imperialism* (London: Dennis Dobson, 1967), *passim*.

¹³ P. A. Baran, *The Political Economy of Growth* (Monthly Review Press, 1957), pp. 44–133.

¹⁴ The proper comparison is between marginal rates of return on investments at home or abroad, not the

Moreover, it is the relatively more buoyant growth of sales and the low wage levels, etc., abroad that are most important in conditioning the expectations about rates of return.

Investments in physical plant and equipment abroad serve to intensify the pressure for and the deeply entrenched quality of imperialism. Thus, the urge to invest capital abroad argues for the worldwide open door to the investor. Private investors have always looked to governmental assistance in opening and keeping open all doors everywhere. So long as commodity trade defined the scope of imperial interests, investing corporations and their governments might safely ignore the policies of foreign governments which did not directly affect their interests. But with the systematic establishment of ownership and management of physical plant and equipment abroad, nearly every detail of a foreign government's policies became a matter of concern.¹⁵

Just as economic development at home affects the relative attractiveness of foreign investment in the capitalist system, so also foreign investments undertaken affect returns to investments in certain domestic lines. As foreign investment grows, this interdependence probably becomes more complex and all-pervasive. While foreign and domestic investment may appear as alternatives to some individual investors, they are mutually interdependent taken together; that is, to some degree, the profitability of domestic investment, and hence employment levels, depends on the profitability of related foreign investments. Thus imperialist control over the environment of foreign investments may well be essential for the profitability of domestic investments.¹⁶

The evidence indicates, for the United States, that marginal rates of return on foreign investments continue to exert a strong attraction for U.S. capital. And the U.S. government is clearly basing its foreign and domestic policies around the expectation of continued major outflows.¹⁷

comparison of marginal rates abroad with average rates at home as in D. K. Fieldhouse, "Imperialism: An Historiographical Revision," *Econ. Hist. Rev.*, Second Series 14 (1961), pp. 187–209. Note also the exception to this generalization in the above discussion of motives for the establishment of foreign subsidiaries.

¹⁵ Cf. L. H. Jenks, *The Migration of British Capital to 1875* (A. A. Knopf, 1927).

¹⁶ Cf. the criticism of Keynes offered in S. H. Frankel, *Some Conceptual Aspects of the International Economic Development of Underdeveloped Territories* (Princeton Univ. Press, 1952), pp. 10–12.

¹⁷ See the interview of M. H. Stans, Secretary of Commerce, in *U.S. News and World Report* (Sept. 8,

IV

The imperialism of modern capitalism—particularly the predominant U.S. variety—consists of a set of economically rational responses by capitalists to the opportunities for economic security and advantage obtainable through the control of foreign economies. The policies pursued by managers (chiefly of the largest oligopolies) toward this objective—supplemented by their bringing state power to bear¹⁸—comprise the pattern of actions of an imperialist capitalism. The fruits of imperialism are some special “gains” from international economic transactions.

On capital account, the story is vivid. From 1950 to 1968 the foreign assets of U.S.-based corporations (i.e., direct investments) rose from about \$12 billion to an estimated \$65 billion, an increase of \$53 billion. In the same period, the net outflow of direct investment from the United States fell short of income received from direct investments by over \$15 billion.¹⁹ These truly impressive gains from U.S. foreign direct investment have involved the shifting of “surplus out of . . . countries and transferring control over a large part of their productive resources to the great imperialist corporations.”²⁰

On trade account the special gains are most clearly seen in terms of the commodity structure of trade. Professor Kenen chooses the phrases “unique characteristics” and “fortunate diversity” in describing the pattern of U.S. exports and imports that has protected the U.S. economy from international commodity price fluctuations and from internally catastrophic import competition.²¹ However, what is “uniquely fortunate” for the U.S. and other major capitalist nations emerged from the past and present control exercised by

these nations over the rest of the world and, to a lesser degree, over one another.

For example, the international economic situation of the British Empire from at least 1870 to 1914 led British corporations and imperial administrators to make all the crucial decisions on the economic development of many empire territories.²² Such decisions—administered usually by political authorities seeking to make the empire a protected economic domain—included the determination of the output mix: agriculture versus mining versus manufacture versus services; cotton versus coffee; and more cotton versus more leisure for some. Such decisions included the size and direction of any investments in both private and public sectors over many decades. Such decisions systematically determined the controlled territories’ tastes, technology, and factor endowments which underlay their international offer curves.

This past pattern within the British Empire is mirrored presently in the other empires, whether they be formal political entities or informal, privately administered, postcolonial spheres of influence; e.g., the U.S. in Latin America. Moreover, this pattern of rich, advanced corporations and their governments historically determining the economic structures of their trading “partners” cannot but call into question the applicability in this instance of the conventional comparative advantage or gains-from-trade analysis of the textbooks which incorrectly assumes the independence of the partners.²³

The point here is that not only are the independence conditions for some gains-from-trade arguments absent, but that a systematically biased development of the international economy is the consequence. It is that bias—as it has operated over the years—that can help to explain the consistently different attitudes of the advanced capitalist nations and the rest of the world toward the performance and promise of the international economy as it is structured today.

V

I have tried to sketch here the basic outlines of some historically developed economic relationships that have led oligopolistic corporations and their governments to pursue more or less coordi-

1969): “In the future it may be necessary for more U.S. companies, in their own interest, to move into the low wage areas of the world and produce for the U.S. market. This is a matter of great concern to us.”

¹⁸ For extensive discussions of this use of state power, see R. Engler, *The Politics of Oil* (Macmillan, 1961), especially pp. 182–229; H. Feis, *Europe the World's Banker, 1870–1914* (Yale, 1930); and S. Nearing and J. Freeman, *Dollar Diplomacy* (Monthly Review Press, 1966).

¹⁹ Various issues of Department of Commerce, *Survey of Current Business*.

²⁰ P. M. Sweezy, “Obstacles to Economic Development,” in C. Feinstein, ed., *Socialism, Capitalism and Economic Growth: Essays Presented to Maurice Dobb* (Cambridge Univ. Press, 1967), p. 195. In *The Age of Imperialism*, Magdoff updates the figures and corroborates the continuing trend.

²¹ P. B. Kenen, *Giant Among Nations* (Rand, McNally, 1964), p. 29.

²² On this point, for a detailed case study of one British colony, see the author's doctoral dissertation, “Economic Aspects of British Colonialism in Kenya, 1895–1930” (Yale Univ., 1969).

²³ Cf. the comparable argument in S. Kuznets, *Economic Growth and Structure* (W. W. Norton, 1965), p. 121.

nated policies aiming to exert control over economic situations in as much of the world as possible. I have tried to suggest that the consequences of such control introduced a bias into the development of the international economy that favored those in control. Lastly it remains to suggest a few examples of the important consequences of imperialism for the allocation of resources within the imperialist economies.

The gains from imperialism have accrued in the first instance to the private owners and managers of capital in the advanced, capitalist countries. In this sense the profits from imperialism remain private. However, the costs of controlling foreign societies are increasingly socialized. Consider the complex economic costs—direct and indirect—of war and the threat of war, of a vast military-dip-

lomatic machine, and of oligopolies whose powers are enhanced through imperialism. Such costs confront us again with the old and fundamental question within the metropolis: who benefits and who loses from the processes of capitalist imperialism—say, for example, from the war in Vietnam?

Perhaps a useful way to conclude is to emphasize two among the necessary directions for further work on the economics of imperialism. We need precise determination of the different groups who gain and lose from imperialism. We also need some careful study of whether and how alternate economic systems in the metropolis can handle their international economic transactions differently.

THE STRUCTURE OF DEPENDENCE*

By THEOTONIO DOS SANTOS
University of Chile

This paper attempts to demonstrate that the dependence of Latin American countries on other countries cannot be overcome without a qualitative change in their internal structures and external relations. We shall attempt to show that the relations of dependence to which these countries are subjected conform to a type of international and internal structure which leads them to underdevelopment or more precisely to a dependent structure that deepens and aggravates the fundamental problems of their peoples.

I. *What is Dependence?*

By dependence we mean a situation in which the economy of certain countries is conditioned by the development and expansion of another economy to which the former is subjected. The relation of interdependence between two or more economies, and between these and world trade, assumes the form of dependence when some countries (the dominant ones) can expand and can be self-sustaining, while other countries (the dependent ones) can do this only as a reflection of that expansion, which can have either a positive or a negative effect on their immediate development [7, p. 6].

The concept of dependence permits us to see the internal situation of these countries as part of world economy. In the Marxian tradition, the theory of imperialism has been developed as a study of the process of expansion of the imperialist centers and of their world domination. In the epoch of the revolutionary movement of the Third World, we have to develop the theory of laws of internal development in those countries that are the object of such expansion and are governed by them. This theoretical step transcends the theory of development which seeks to explain the situation of the underdeveloped countries as a product of their slowness or failure to adopt the patterns of efficiency characteristic of developed

countries (or to "modernize" or "develop" themselves). Although capitalist development theory admits the existence of an "external" dependence, it is unable to perceive underdevelopment in the way our present theory perceives it, as a consequence and part of the process of the world expansion of capitalism—a part that is necessary to and integrally linked with it.

In analyzing the process of constituting a world economy that integrates the so-called "national economies" in a world market of commodities, capital, and even of labor power, we see that the relations produced by this market are unequal and combined—unequal because development of parts of the system occurs at the expense of other parts. Trade relations are based on monopolistic control of the market, which leads to the transfer of surplus generated in the dependent countries to the dominant countries; financial relations are, from the viewpoint of the dominant powers, based on loans and the export of capital, which permit them to receive interest and profits; thus increasing their domestic surplus and strengthening their control over the economies of the other countries. For the dependent countries these relations represent an export of profits and interest which carries off part of the surplus generated domestically and leads to a loss of control over their productive resources. In order to permit these disadvantageous relations, the dependent countries must generate large surpluses, not in such a way as to create higher levels of technology but rather superexploited manpower. The result is to limit the development of their internal market and their technical and cultural capacity, as well as the moral and physical health of their people. We call this combined development because it is the combination of these inequalities and the transfer of resources from the most backward and dependent sectors to the most advanced and dominant ones which explains the inequality, deepens it, and transforms it into a necessary and structural element of the world economy.

II. *Historic Forms of Dependence*

Historic forms of dependence are conditioned by: (1) the basic forms of this world economy which has its own laws of development; (2) the type of economic relations dominant in the capi-

* This work expands on certain preliminary work done in a research project on the relations of dependence in Latin America, directed by the author at the Center for Socio-Economic Studies of the Faculty of Economic Science of the University of Chile. In order to abridge the discussion of various aspects, the author was obliged to cite certain of his earlier works. The author expresses his gratitude to the researcher Orlando Caputo and Roberto Pizarro for some of the data utilized and to Sergio Ramos for his critical comments on the paper.

talist centers and the ways in which the latter expand outward; and (3) the types of economic relations existing inside the peripheral countries which are incorporated into the situation of dependence within the network of international economic relations generated by capitalist expansion. It is not within the purview of this paper to study these forms in detail but only to distinguish broad characteristics of development.

Drawing on an earlier study, we may distinguish: (1) Colonial dependence, trade export in nature, in which commercial and financial capital in alliance with the colonialist state dominated the economic relations of the Europeans and the colonies, by means of a trade monopoly complemented by a colonial monopoly of land, mines, and manpower (serf or slave) in the colonized countries. (2) Financial-industrial dependence which consolidated itself at the end of the nineteenth century, characterized by the domination of big capital in the hegemonic centers, and its expansion abroad through investment in the production of raw materials and agricultural products for consumption in the hegemonic centers. A productive structure grew up in the dependent countries devoted to the export of these products (which Levin labeled export economies [11]; other analysis in other regions [12] [13]), producing what ECLA has called "foreign-oriented development" (*desarrollo hacia afuera*) [4]. (3) In the postwar period a new type of dependence has been consolidated, based on multinational corporations which began to invest in industries geared to the internal market of underdeveloped countries. This form of dependence is basically technological-industrial dependence [6].

Each of these forms of dependence corresponds to a situation which conditioned not only the international relations of these countries but also their internal structures: the orientation of production, the forms of capital accumulation, the reproduction of the economy, and, simultaneously, their social and political structure.

III. The Export Economies

In forms (1) and (2) of dependence, production is geared to those products destined for export (gold, silver, and tropical products in the colonial epoch; raw materials and agricultural products in the epoch of industrial-financial dependence); i.e., production is determined by demand from the hegemonic centers. The internal productive structure is characterized by rigid specialization and monoculture in entire regions (the Caribbean, the Brazilian Northeast, etc.). Alongside these export sectors there grew up certain com-

plementary economic activities (cattle-raising and some manufacturing, for example) which were dependent, in general, on the export sector to which they sell their products. There was a third, subsistence economy which provided manpower for the export sector under favorable conditions and toward which excess population shifted during periods unfavorable to international trade.

Under these conditions, the existing internal market was restricted by four factors: (1) Most of the national income was derived from export, which was used to purchase the inputs required by export production (slaves, for example) or luxury goods consumed by the hacienda- and mine-owners, and by the more prosperous employees. (2) The available manpower was subject to very arduous forms of superexploitation, which limited its consumption. (3) Part of the consumption of these workers was provided by the subsistence economy, which served as a complement to their income and as a refuge during periods of depression. (4) A fourth factor was to be found in those countries in which land and mines were in the hands of foreigners (cases of an enclave economy): a great part of the accumulated surplus was destined to be sent abroad in the form of profits, limiting not only internal consumption but also possibilities of reinvestment [1]. In the case of enclave economies the relations of the foreign companies with the hegemonic center were even more exploitative and were complemented by the fact that purchases by the enclave were made directly abroad.

IV. The New Dependence

The new form of dependence, (3) above, is in process of developing and is conditioned by the exigencies of the international commodity and capital markets. The possibility of generating new investments depends on the existence of financial resources in foreign currency for the purchase of machinery and processed raw materials not produced domestically. Such purchases are subject to two limitations: the limit of resources generated by the export sector (reflected in the balance of payments, which includes not only trade but also service relations); and the limitations of monopoly on patents which leads monopolistic firms to prefer to transfer their machines in the form of capital rather than as commodities for sale. It is necessary to analyze these relations of dependence if we are to understand the fundamental structural limits they place on the development of these economies.

1. Industrial development is dependent on an export sector for the foreign currency to buy the

inputs utilized by the industrial sector. The first consequence of this dependence is the need to preserve the traditional export sector, which limits economically the development of the internal market by the conservation of backward relations of production and signifies, politically, the maintenance of power by traditional decadent oligarchies. In the countries where these sectors are controlled by foreign capital, it signifies the remittance abroad of high profits, and political dependence on those interests. Only in rare instances does foreign capital not control at least the marketing of these products. In response to these limitations, dependent countries in the 1930's and 1940's developed a policy of exchange restrictions and taxes on the national and foreign export sector; today they tend toward the gradual nationalization of production and toward the imposition of certain timid limitations on foreign control of the marketing of exported products. Furthermore, they seek, still somewhat timidly, to obtain better terms for the sale of their products. In recent decades, they have created mechanisms for international price agreements, and today UNCTAD and ECLA press to obtain more favorable tariff conditions for these products on the part of the hegemonic centers. It is important to point out that the industrial development of these countries is dependent on the situation of the export sector, the continued existence of which they are obliged to accept.

2. Industrial development is, then, strongly conditioned by fluctuations in the balance of payments. This leads toward deficit due to the relations of dependence themselves. The causes of the deficit are three:

a) Trade relations take place in a highly monopolized international market, which tends to lower the price of raw materials and to raise the prices of industrial products, particularly inputs. In the second place, there is a tendency in modern technology to replace various primary products with synthetic raw materials. Consequently the balance of trade in these countries tends to be less favorable (even though they show a general surplus). The overall Latin American balance of trade from 1946 to 1968 shows a surplus for each of those years. The same thing happens in almost every underdeveloped country. However, the losses due to deterioration of the terms of trade (on the basis of data from ECLA and the International Monetary Fund), excluding Cuba, were \$26,383 million for the 1951-66 period, taking 1950 prices as a base. If Cuba and Venezuela are excluded, the total is \$15,925 million.

b) For the reasons already given, foreign capi-

tal retains control over the most dynamic sectors of the economy and repatriates a high volume of profit; consequently, capital accounts are highly unfavorable to dependent countries. The data show that the amount of capital leaving the country is much greater than the amount entering; this produces an enslaving deficit in capital accounts. To this must be added the deficit in certain services which are virtually under total foreign control—such as freight transport, royalty payments, technical aid, etc. Consequently, an important deficit is produced in the total balance of payments; thus limiting the possibility of importation of inputs for industrialization.

c) The result is that "foreign financing" becomes necessary, in two forms: to cover the existing deficit, and to "finance" development by means of loans for the stimulation of investments and to "supply" an internal economic surplus which was decapitalized to a large extent by the remittance of part of the surplus generated domestically and sent abroad as profits.

Foreign capital and foreign "aid" thus fill up the holes that they themselves created. The real value of this aid, however, is doubtful. If overcharges resulting from the restrictive terms of the aid are subtracted from the total amount of the grants, the average net flow, according to calculations of the Inter-American Economic and Social Council, is approximately 54 percent of the gross flow [5].

If we take account of certain further facts—that a high proportion of aid is paid in local currencies, that Latin American countries make contributions to international financial institutions, and that credits are often "tied"—we find a "real component of foreign aid" of 42.2 percent on a very favorable hypothesis and of 38.3 percent on a more realistic one [5, II-33]. The gravity of the situation becomes even clearer if we consider that these credits are used in large part to finance North American investments, to subsidize foreign imports which compete with national products, to introduce technology not adapted to the needs of underdeveloped countries, and to invest in low-priority sectors of the national economies. The hard truth is that the underdeveloped countries have to pay for all of the "aid" they receive. This situation is generating an enormous protest movement by Latin American governments seeking at least partial relief from such negative relations.

3. Finally, industrial development is strongly conditioned by the technological monopoly exercised by imperialist centers. We have seen that the underdeveloped countries depend on the importation of machinery and raw materials for the

development of their industries. However, these goods are not freely available in the international market; they are patented and usually belong to the big companies. The big companies do not sell machinery and processed raw materials as simple merchandise: they demand either the payment of royalties, etc., for their utilization or, in most cases, they convert these goods into capital and introduce them in the form of their own investments. This is how machinery which is replaced in the hegemonic centers by more advanced technology is sent to dependent countries as capital for the installation of affiliates. Let us pause and examine these relations, in order to understand their oppressive and exploitative character.

The dependent countries do not have sufficient foreign currency, for the reasons given. Local businessmen have financing difficulties, and they must pay for the utilization of certain patented techniques. These factors oblige the national bourgeois governments to facilitate the entry of foreign capital in order to supply the restricted national market, which is strongly protected by high tariffs in order to promote industrialization. Thus, foreign capital enters with all the advantages: in many cases, it is given exemption from exchange controls for the importation of machinery; financing of sites for installation of industries is provided; government financing agencies facilitate industrialization; loans are available from foreign and domestic banks, which prefer such clients; foreign aid often subsidizes such investments and finances complementary public investments; after installation, high profits obtained in such favorable circumstances can be reinvested freely. Thus it is not surprising that the data of the U.S. Department of Commerce reveal that the percentage of capital brought in from abroad by these companies is but a part of the total amount of invested capital. These data show that in the period from 1946 to 1967 the new entries of capital into Latin America for direct investment amounted to \$5,415 million, while the sum of reinvested profits was \$4,424 million. On the other hand, the transfers of profits from Latin America to the United States amounted to \$14,775 million. If we estimate total profits as approximately equal to transfers plus reinvestments we have the sum of \$18,983 million. In spite of enormous transfers of profits to the United States, the book value of the United States's direct investment in Latin America went from \$3,045 million in 1946 to \$10,213 million in 1967. From these data it is clear that: (1) Of the new investments made by U.S. companies in Latin America for the period 1946-67, 55 percent corre-

sponds to new entries of capital and 45 percent to reinvestment of profits; in recent years, the trend is more marked, with reinvestments between 1960 and 1966 representing more than 60 percent of new investments. (2) Remittances remained at about 10 percent of book value throughout the period. (3) The ratio of remitted capital to new flow is around 2.7 for the period 1946-67; that is, for each dollar that enters \$2.70 leaves. In the 1960's this ratio roughly doubled, and in some years was considerably higher.

The *Survey of Current Business* data on sources and uses of funds for direct North American investment in Latin America in the period 1957-64 show that, of the total sources of direct investment in Latin America, only 11.8 percent came from the United States. The remainder is in large part, the result of the activities of North American firms in Latin America (46.4 percent net income, 27.7 percent under the heading of depreciation), and from "sources located abroad" (14.1 percent). It is significant that the funds obtained abroad that are external to the companies are greater than the funds originating in the United States.

V. Effects on the Productive Structure

It is easy to grasp, even if only superficially, the effects that this dependent structure has on the productive system itself in these countries and the role of this structure in determining a specified type of development, characterized by its dependent nature.

The productive system in the underdeveloped countries is essentially determined by these international relations. In the first place, the need to conserve the agrarian or mining export structure generates a combination between more advanced economic centers that extract surplus value from the more backward sectors, and also between internal "metropolitan" centers and internal interdependent "colonial" centers [10]. The unequal and combined character of capitalist development at the international level is reproduced internally in an acute form. In the second place the industrial and technological structure responds more closely to the interests of the multinational corporations than to internal developmental needs (conceived of not only in terms of the overall interests of the population, but also from the point of view of the interests of a national capitalist development). In the third place, the same technological and economic-financial concentration of the hegemonic economies is transferred without substantial alteration to very different economies and societies, giving rise to a highly unequal pro-

ductive structure, a high concentration of incomes, underutilization of installed capacity, intensive exploitation of existing markets concentrated in large cities, etc.

The accumulation of capital in such circumstances assumes its own characteristics. In the first place, it is characterized by profound differences among domestic wage-levels, in the context of a local cheap labor market, combined with a capital-intensive technology. The result, from the point of view of relative surplus value, is a high rate of exploitation of labor power. (On measurements of forms of exploitation, see [3].)

This exploitation is further aggravated by the high prices of industrial products enforced by protectionism, exemptions and subsidies given by the national governments, and "aid" from hegemonic centers. Furthermore, since dependent accumulation is necessarily tied into the international economy, it is profoundly conditioned by the unequal and combined character of international capitalist economic relations, by the technological and financial control of the imperialist centers by the realities of the balance of payments, by the economic policies of the state, etc. The role of the state in the growth of national and foreign capital merits a much fuller analysis than can be made here.

Using the analysis offered here as a point of departure, it is possible to understand the limits that this productive system imposes on the growth of the internal markets of these countries. The survival of traditional relations in the countryside is a serious limitation on the size of the market, since industrialization does not offer hopeful prospects. The productive structure created by dependent industrialization limits the growth of the internal market.

First, it subjects the labor force to highly exploitative relations which limit its purchasing power. Second, in adopting a technology of intensive capital use, it creates very few jobs in comparison with population growth, and limits the generation of new sources of income. These two limitations affect the growth of the consumer goods market. Third, the remittance abroad of profits carries away part of the economic surplus generated within the country. In all these ways limits are put on the possible creation of basic national industries which could provide a market for the capital goods this surplus would make possible if it were not remitted abroad.

From this cursory analysis we see that the alleged backwardness of these economies is not due to a lack of integration with capitalism but that, on the contrary, the most powerful obstacles to

their full development come from the way in which they are joined to this international system and its laws of development.

VI. *Some Conclusions: Dependent Reproduction*

In order to understand the system of dependent reproduction and the socioeconomic institutions created by it, we must see it as part of a system of world economic relations based on monopolistic control of large-scale capital, on control of certain economic and financial centers over others, on a monopoly of a complex technology that leads to unequal and combined development at a national and international level. Attempts to analyze backwardness as a failure to assimilate more advanced models of production or to modernize are nothing more than ideology disguised as science. The same is true of the attempts to analyze this international economy in terms of relations among elements in free competition, such as the theory of comparative costs which seeks to justify the inequalities of the world economic system and to conceal the relations of exploitation on which it is based [14].

In reality we can understand what is happening in the underdeveloped countries only when we see that they develop within the framework of a process of dependent production and reproduction. This system is a dependent one because it reproduces a productive system whose development is limited by those world relations which necessarily lead to the development of only certain economic sectors, to trade under unequal conditions [9], to domestic competition with international capital under unequal conditions, to the imposition of relations of superexploitation of the domestic labor force with a view to dividing the economic surplus thus generated between internal and external forces of domination. (On economic surplus and its utilization in the dependent countries, see [1].)

In reproducing such a productive system and such international relations, the development of dependent capitalism reproduces the factors that prevent it from reaching a nationally and internationally advantageous situation; and it thus reproduces backwardness, misery, and social marginalization within its borders. The development that it produces benefits very narrow sectors, encounters unyielding domestic obstacles to its continued economic growth (with respect to both internal and foreign markets), and leads to the progressive accumulation of balance-of-payments deficits, which in turn generate more dependence and more superexploitation.

The political measures proposed by the devel-

opmentalists of ECLA, UNCTAD, BID, etc., do not appear to permit destruction of these terrible chains imposed by dependent development. We have examined the alternative forms of development presented for Latin America and the dependent countries under such conditions elsewhere [8]. Everything now indicates that what can be expected is a long process of sharp political and military confrontations and of profound social radicalization which will lead these countries to a dilemma: governments of force which open the way to facism, or popular revolutionary governments, which open the way to socialism. Intermediate solutions have proved to be, in such a contradictory reality, empty and utopian.

REFERENCES

1. PAUL BARAN, *Political Economy of Growth* (Monthly Review Press, 1967).
2. THOMAS BALOGH, *Unequal Partners* (Basil Blackwell, 1963).
3. PABLO GONZALEZ CASANOVA, *Sociología de la explotación, Siglo XXI* (México, 1969).
4. CEPAL, *La CEPAL y el Análisis del Desarrollo Latinoamericano* (1968, Santiago, Chile).
5. Consejo Interamericano Económico Social (CIES) O.A.S., Interamerican Economic and Social Council, External Financing for Development in L.A. *El Financiamiento Externo para el Desarrollo de América Latina* (Pan-American Union, Washington, 1969).
6. THEOTONIO DOS SANTOS, *El nuevo carácter de la dependencia*, CESO (Santiago de Chile, 1968).
7. ———, *La crisis de la teoría del desarrollo y las relaciones de dependencia en América Latina*, Boletín del CESO, 3 (Santiago, Chile, 1968).
8. ———, *La dependencia económica y las alternativas de cambio en América Latina*, Ponencia al IX Congreso Latinoamericano de Sociología (México, Nov., 1969).
9. A. EMMANUEL, *L'Echange Inégal* (Maspero, Paris, 1969).
10. ANDRE G. FRANK, *Development and Underdevelopment in Latin America* (Monthly Review Press, 1968).
11. I. V. LEVIN, *The Export Economies* (Harvard Univ. Press, 1964).
12. GUNNAR MYRDAL, *Asian Drama* (Pantheon, 1968).
13. K. NKRUMAH, *Neocolonialismo, última etapa del imperialismo, Siglo XXI* (México, 1966).
14. CRISTIAN PALLOIX, *Problemes de la Croissance en Economie Ouverte* (Maspero, Paris, 1969).

MILITARISM AND IMPERIALISM

By HARRY MAGDOFF
New School for Social Research

Peace reigns supreme in the realm of neoclassical economics. War, militarism, and the pacification of natives are treated as merely elements which disturb the harmonious equilibrium models which are to supply us with the universal truths about the allocation of scarce resources.

One of the distinguishing features of Marxist thought, on the other hand, is the conviction that economic processes must be understood as part of a social organism in which political force plays a leading role and in which war is at least as typical as peace. In this context, militarism and imperialism are seen as major determinants of the form and direction of technological change, of the allocation of resources within a country, and of the allocation of resources between countries (notably between rich and poor countries). Accordingly, price and income relations, treated as the ultimate yardsticks of economic efficiency and social justice in neoclassical economics, are viewed, in the Marxist context, as evolutionary products of capitalist institutions in which political force and "pure" economics are intertwined. Rosa Luxemburg put the Marxist case this way:

Bourgeois liberal theory takes into account only [one aspect of economic development]: the realm of "peaceful competition," the marvels of technology and pure commodity exchange; it separates it strictly from the other aspect: the realm of capital's blustering violence which is regarded as more or less incidental to foreign policy and quite independent of the economic sphere of capital.

In reality, political power is nothing but a vehicle for the economic process. The conditions for the reproduction of capital provide the organic link between these two aspects of the accumulation of capital. The historical career of capitalism can be appreciated only by taking them together.¹

The facts of U.S. history provide eloquent testimony to the accuracy of this diagnosis. Thus, Professor Quincy Wright, who directed a major study of war under the auspices of the University of Chicago, observed in 1942:

The United States, which has, perhaps somewhat unjustifiably, prided itself on its peacefulness, has had only twenty years during its entire history when its army or navy has not been in active operation some days, somewhere.²

¹ Rosa Luxemburg, *The Accumulation of Capital* (New York, 1964), pp. 452-53.

² Quincy Wright, *A Study of War* (Chicago, 1942), Vol. I, p. 236.

Professor Wright identifies years of peace as those in which no action of any sort occurred. A more revealing picture is obtained if we measure months of war against months of peace and bring the information up to the present. Adding up the months during which U.S. military forces were engaged in action—starting from the Revolutionary War and including wars against the Indians, punitive expeditions to Latin America and Asia, as well as major wars—we find that the United States was engaged in warlike activity during three-fourths of its history, in 1,782 of the last 2,340 months.³ In other words, on the average, there have been three full years in which our armed forces have been engaged in action for every full year of peace. This comparison does not indicate the full extent of the use of military power by the United States to enforce its will. For example, it does not include activities such as those formerly conducted by U.S. gunboats in a "constant patrol in the Yangtze River . . . from the mouth of the river up nearly 2,000 miles into the very heart of China."⁴

It should therefore come as no surprise to discover that war-related expenditures have constituted the dominant sector of the federal budget throughout our history. Omitting the years of the second World War and the postwar period, where the record is so well known, a tabulation of federal expenditures by decade, from 1800 to 1939, for army, navy, veterans' compensation and pensions, and interest on the debt—prior to the New Deal federal debt incurred was primarily a result of war spending—shows that except for one decade, at least 54 percent of federal expenditures were for military activities or preparations during the decade or to meet obligations arising from previous military activity.⁵ The one exception was the decade of the Great Depression (1930-39) when the percentage dropped to somewhat below 40 percent. In seven of the fourteen decades the war-related share of the federal budget was 70 percent or more.

³ Calculated from list in Lawrence Dennis, *Operational Thinking for Survival* (Colorado Springs, 1969), Appendix II.

⁴ Office of Naval Intelligence, *The United States Navy as an Industrial Asset* (Washington, D.C., 1923), p. 4.

⁵ Calculated from data in *Historical Statistics of the United States, Colonial Times to 1957* (Washington, D.C., 1961), pp. 718-19.

This almost continuous preoccupation with military affairs was clearly not inspired by fears of invading barbarians. Of course, the competing colonial and commercial interests of France, England, Spain, and Russia were part of the reality in which the infant and adolescent United States had to operate. At times, self-defense had to be considered. Moreover, resolution of internal tensions, as in the Civil War, exercised a major influence on military aspects of U.S. life. All of this, however, occurred within a context of empire building, for there has been a continuous thread in U.S. history, beginning with colonial and revolutionary days, of economic, political, and military expansionism directed towards the creation and growth of an American empire. The original expansionism, for which military investment was needed, concentrated on three main thrusts: (1) consolidation of a transcontinental nation, (2) obtaining control of the Caribbean area, and (3) achieving a major position in the Pacific Ocean.⁶ It should be noted that this expansionism was not confined to what is now considered the continental territory of the United States: striving for control of the seas, as a shield and promoter of international commerce, has been an ingredient of U.S. policy from its earliest days. In fact, the struggle to incorporate the West Coast into the United States was, among other things, prompted by the desire to control Pacific Ocean ports for the Asian trade.⁷

The experience thus gained in the early stages of empire building turned out to be most useful when the leading nations of the world entered the stage of imperialism. Several decisive and coinciding developments in the late nineteenth and early twentieth centuries mark off this new stage:

1. The onset of significant concentration of economic power in the hands of a relatively small number of industrial and financial giants in advanced nations. Competing interest-groups continued to exist, but now the success or failure of the advanced economies became closely identified with the prosperity of the new giant corporations whose *modus operandi* required control over international sources of supply and markets.

2. The decline of Great Britain's monopoly position as world trader and world banker. The burgeoning competitive industrial powers—notably, Germany, France, the United States, and Japan—pressed for a reshuffle of established trade relations and a redistribution of world markets.

3. Industrialization and new naval technology

⁶ Richard W. Van Alstyne, *The Rising American Empire* (Chicago, 1965).

⁷ *Ibid.*, Chap. 5 ("Manifest Destiny and Empire, 1820-1870").

enabled competitive nations to build up their own naval strength to the point where Great Britain could no longer maintain unilateral control over the major sea lanes. As Quincy Wright put it in the study already referred to, "naval inventions and the spread of industrialization had ended the *pax Britannica*."⁸ Control over sea routes also involved establishing military bases where naval units could be refueled and repaired. The availability of decisive mobile military power on the one hand required acquisition of strategic foreign territory to support bases and on the other hand provided the means for aggressive pursuit of colonial possessions.

4. The earliest stage of the new imperialism engendered a race by the major powers for control of available foreign real estate. According to Theodore Ropp, after 1880 "every great power except Austria-Hungary . . . became involved in . . . active, conscious colonial expansionism. . . ."⁹ Of the traditional colonial powers—the Netherlands, Portugal, Spain, Britain, France, and Russia—the last four continued to add to their holdings. (Spain, after losing Cuba and the Philippines, proceeded to conquer Spanish Morocco.) And at the same time five new powers entered the race for colonial territory: Germany, Italy, Belgium, Japan, and the United States. As for the United States, it was the Spanish-American War, of course, that placed it with both feet in the imperialist camp. And it was success in this war, plus the subsequent pacification of the Cuban and Philippine natives, which satisfied two long-term U.S. expansionist ambitions: a leading position in the Caribbean, broadening the highway to the rest of Latin America, and a solid base in the Pacific for a greater stake in Asian business.

As far as the United States is concerned, there have been three distinct stages in the drive to empire: (1) the period when the United States was the supplier of food and raw materials to the rest of the world, when it was an importer of capital, and when maritime commercial interests were relatively very strong; (2) the period when the United States began to compete with other industrialized nations as an exporter of manufactured goods and an exporter of capital—a time when a small number of industrial and financial giants began to dominate the economic scene; and (3) the period when the United States becomes the major, dominant capitalist economy, the largest manufacturer, foreign investor, trader, the world's

⁸ Quincy Wright, *op. cit.*, Vol. I, p. 299.

⁹ Theodore Ropp, *War in the Modern World* (New York, 1962), p. 206.

banker and the dollar becomes the key international currency.

The energy and determination with which the expansionist strategy is pursued change from time to time. In the transition from one period to another, and because of internal as well as external conditions, it appears at times as if the United States is "isolationist" and uninterested in further extension of its influence and control.¹⁰ Yet it is especially noteworthy that the drive for business opportunities on a world scale is ever present. Even when, as in New Deal days, domestic solutions were sought for crises, the development of foreign business was high on the agenda of government and private enterprise. Given the structure of the economy, the major operating levers work in such a way as to repeatedly reassert expansionism as the dominant strategy. In this perspective, the history of the years since the end of the second World War are far from a new departure; instead, they are a culmination of long-term tendencies which profited by and matured most readily in the environment created by the course of the last major war.

The postwar leap forward in empire building and the transition of U.S. society to rampant militarism are associated with two phenomena: (1) the desire to resist and repress socialist nations and to defeat national liberation movements designed to release underdeveloped countries from dependence on the imperialist network and (2) the extension of U.S. power to fill "vacuums" created by the decline of Western European and Japanese influence in Asia, Africa, and Latin America.

Combating the rise of socialism is of course not a new objective. The destruction of the Russian Revolution was a top priority of the imperialist powers beginning in 1917. In this connection, Thorstein Veblen's observations on the Versailles Treaty in his 1920 review of Keynes's *The Economic Consequences of the Peace* are most pertinent:

The events of the past months go to show that the central and most binding provision of the Treaty (and of the League) is an unrecorded clause by which the governments of the Great Powers are banded together for the suppression of Soviet Russia—unrecorded unless record of it is to be found somewhere among the secret archives of the League or of the Great Powers. Apart from this unacknowledged compact there appears to be nothing in the Treaty that

has any character of stability or binding force. Of course, this compact for the reduction of Soviet Russia was not written into the text of the Treaty; it may rather be said to have been the parchment upon which the text was written.¹¹

The failure of the United States to join the League of Nations reflected no slackness in its efforts to contain anti-imperialist revolutions: in Russia, these efforts took the form of armed intervention and support of anti-Bolshevik forces with food and other economic supplies; in Hungary, the manipulation of food supplies to help defeat the Bela Kun government. Surely the issue at that time was not fear of aggressive Russian or Hungarian militarism. Nor can much credit be given to political or religious idealism. The relevant motive, clearly, was recovery of territory lost to free enterprise and prevention of the spread of the contagious revolutionary disease to Western Europe and the colonies. Any such spread, it was recognized, would severely affect the stability and prosperity of the remaining capitalist nations.

Capitalism as an economic system was never confined to one nation. It was born, developed, and prospered as part of a world system. Karl Marx went so far as to claim "the specific task of bourgeois society is the establishment of a world market, at least in outline, and of production based upon this world market."¹² One might add that it has been the specific task of imperialism to fill out this outline and establish a complex international network of trade, finance, and investment. Given this network, it follows that limitation of opportunity to trade and invest in one part of the world restricts to a greater or lesser extent the freedom of action of private enterprise in other parts of the world. The dimensions of the defense of free enterprise therefore become worldwide.

The United States had long ago accepted its destiny to open and keep open the door for trade and investment in other parts of the world. The obstacles were not only the heathens who wanted to be left alone, but the preference systems established in the colonies of the older nations. The decline of political colonialism and the weakness of the other great powers thus placed upon the United States a primary responsibility for the defense of the capitalist system and at the same

¹⁰ Thorstein Veblen, "The Economic Consequences of the Peace," in *Essay in Our Changing Order* (New York, 1934), p. 464.

¹¹ In a letter from Marx to Engels, Oct. 8, 1858, in Karl Marx and Friedrich Engels, *Correspondence 1846-1895* (New York, 1934), p. 117.

¹² The isolationism was usually more apparent than real. See William Appleman Williams, *The Tragedy of American Diplomacy* (New York, 1962), Chap. 4, "The Legend of Isolationism."

time afforded golden opportunities to obtain special beachheads and open doors for U.S. enterprise.

With a task of this magnitude, it is little wonder that the United States now has a larger "peacetime" war machine, covering a greater part of the globe, than has any other nation in all of past history. Imperialism necessarily involves militarism. Indeed, they are twins that have fed on each other in the past, as they do now. Yet not even at the peak of the struggle for colonies did any of the imperialist powers, or combination of powers, maintain a war machine of such size and such dispersion as does the United States today. In 1937, when the arms race in preparation for the second World War was already under way, the per capita military expenditures of all the great powers combined—the United States, the British Empire, France, Japan, Germany, Italy, and the Soviet Union—was \$25. (Germany's per capita of \$58.82 was then the largest.)¹³ In 1968, the per capita military expenditures of the United States alone, in 1937 prices, was \$132. This was only in part due to the Vietnam war: in 1964, our most recent "peace" year, the per capita military expenditures in 1937 prices was \$103.¹⁴

One of the reasons for this huge increase in military outlays is no doubt the greater sophistication of weaponry. (By the same token, it is the advanced airplane and missile technology which makes feasible the U.S. globe-straddling military posture.) An additional reason, of course, is the military strength of the socialist camp. I would like to suggest a third reason: that a substantial portion of the huge military machine, including that of the Western European nations, is the price being paid to maintain the imperialist network of trade and investment in the absence of colonialism. The achievement of political independence by former colonies has stimulated internal class struggles in the new states for economic as well as political independence. Continuing the economic dependence of these nations on the metropolitan centers within the framework of political independence calls for, among other things, the worldwide dispersion of U.S. military forces

and the direct military support of the local ruling classes.

Precise information on the dispersion of U.S. forces is kept an official secret. However, retired General David M. Shoup, former head of the Marine Corps, who should be in a position to make a realistic estimate, stated in a recent article in *The Atlantic*: "We maintain more than 1,517,000 Americans in uniform overseas in 119 countries. We have 8 treaties to help defend 48 nations if they ask us to or if we choose to intervene in their affairs."¹⁵ The main substance of U.S. overseas power, aside from its present application in Vietnam, is spread out over 429 major and 2,972 minor military bases. These bases cover 4,000 square miles in 30 foreign countries, as well as Hawaii and Alaska.¹⁶ Backing this up, and acting as a coordinator of the lesser imperialist powers and the Third World incorporated in the imperialist network, is a massive program of military assistance. According to a recent study:

U.S. military aid . . . since 1945 has averaged more than \$2 billion per year. It rose to as much as \$5 billion in fiscal year (FY) 1952 and fell to as low as \$831 million in FY 1956. The number of recipient countries rose from 14 in 1950 to a peak so far of 69 in 1963. In all, some 80 countries have received a total of \$50 billion in American military aid since World War II. Except for 11 hard-core communist countries and certain nations tied closely to either Britain or France, very few nations have never received military aid of one kind or another from the United States.¹⁷

The above factual recital by no means exhausts the international functions of U.S. militarism. Space considerations permit no more than passing reference to (a) the active promotion of commercial armament sales abroad (contributing a sizable portion of the merchandise export surplus in recent years), (b) the extensive training of foreign military personnel, and (c) the use of economic-aid funds to train local police forces for "handling mob demonstrations and counterintelligence work."¹⁸ These are, in the main, addi-

¹³ Quincy Wright, *op. cit.*, pp. 670–71.

¹⁴ The data on military expenditures are the purchases of goods and services for "national defense" and "space research and technology" as used in computing gross national product. The 1964 and 1968 data are reported in the *Survey of Current Business*, July, 1968, and July, 1969. The adjustment for price changes was made by using the implicit price deflators for federal purchases of goods and services, as given in the *Economic Report of the President*, Jan., 1969.

¹⁵ General David M. Shoup, "The New American Militarism," *The Atlantic*, Apr., 1969. The figure of 119 countries seems too large. General Shoup was probably including bases on island locations, which he counted as separate countries. Our guess is that U.S. armed forces to man bases, administer military assistance, and train foreign officers are located in 70 to 80 countries.

¹⁶ *New York Times*, Apr. 9, 1969.

¹⁷ George Thayer, *The War Business, The International Trade in Armaments* (New York, 1969), pp. 37–38. This is a summary of data presented in *Military Assistance Facts*, May 1, 1966, brought up to date through fiscal year 1968.

¹⁸ For (a) see *ibid.* (b) See John Dunn, "Military

tional instruments for maintaining adherence and loyalty of the nonsocialist world to the free enterprise system in general, and to the United States in particular.

The military forces of the politically independent underdeveloped countries frequently perform a very special function. This arises from the relative weaknesses of the competitive elite power groups: large landowners, merchants, industrialists and financiers—each with varying degrees of alliance to interest groups in the metropolitan center. When none of these ruling-class groups has the strength and resources to take the political reins in its own hands and assert its hegemony over the others, the social order is operated by means of temporary and unstable alliances. Under such circumstances, and especially when the existing order is threatened by social revolution, the military organizations become increasingly important as a focal point for the power struggle within the ruling classes and/or as the organizer of political arrangements. Space limitations do not permit a review of this special role of militarism in the underdeveloped world as, one might say, the skeletal framework of the imperialist system in the absence of colonies. It is this framework that is supported and nurtured by the practices mentioned above: military training and advisory services, the widespread military assistance programs, and the stimulus given to commercial sales of U.S. armaments.

This militarism which is working to control the rest of the world is at the same time helping to shape the nature of U.S. society. Some sense of the immensity of this impact can be obtained by noting the relevance of military spending on the employment/unemployment situation. In the first three quarters of 1969, approximately 8.3 million persons were employed as a result of the military program: 3.5 million in the armed services, 1.3 million Defense Department civilian employees, and 3.5 million engaged in producing and moving the goods purchased for the military services.¹⁹

Aid and Military Elites: The Political Potential of American Training and Technical Assistance Programs" (unpublished Ph.D. dissertation, Princeton Univ., 1961). (c) See Edwin Lieuwen, *The United States and the Challenge to Security in Latin America* (Ohio, 1966), p. 16.

¹⁹ Data on armed services and Defense Department civilian employment from *Defense Indicators* (Bureau of the Census), Nov., 1969. The estimate of the number employed by private industry for military production is based on Richard P. Oliver's study, "The Employment Effect of Defense Expenditures," *Monthly Labor Rev.*, Sept., 1967. Mr. Oliver estimated 2,972 million employed in private industry in the fiscal year

At the same time, there are at least 3.7 million unemployed.²⁰

Consider for a moment what it would mean if 8.3 million were not engaged in military affairs. Without substitute employment, this could mean a total of over 12 million unemployed, or a 14.3 percent rate of unemployment. The last time the United States had such a rate of unemployment was 1937. The percentage of the labor force unemployed in 1931, the second full year of the Great Depression, was less than 2 points higher, 15.9 percent.²¹

So far we have not taken into account the multiplier effect. It has been estimated that for every \$1.00 spent on national defense, another \$1.00 to \$1.40 of national product is stimulated.²² If we accept only the lower estimate and assume for the sake of the argument equivalent labor productivity in the military and civilian sectors, we reach a measure of unemployment in the neighborhood of 24.3 percent, in the absence of the military budget. Compare this with the unemployment rate of 24.9 percent at the depth of the depression in 1932.

A counterargument can, of course, be made to these broad generalizations. Unemployment insurance, for example, would to a limited extent and for a very limited time act as an offset. Conceivably, a sharp decline in military spending, if there were no financial collapse accompanying it, would reduce interest rates and thus perhaps stimulate construction and some types of state and municipal investment. A reduction in taxes would generate consumer demand. A rise in the federal social

ending June 30, 1967, as a result of Defense Department expenditures. We brought this estimate up to date by (1) assuming no increase in productivity or major change in the composition of production since fiscal year 1967; (2) using the expenditure data for the first three quarters of 1969; (3) adding space research and technology and one-half of Atomic Energy Commission expenditures, both of which had been excluded in Mr. Oliver's estimates; and (4) adjusting for price increases in the last two years. The resulting figure of 3.5 million is therefore a broad estimate, but the margin of error is not such as in any way to invalidate our analysis.

²⁰ Based on data in *Employment and Earnings* (Bureau of Labor Statistics), Jan. to Nov., 1969. The 3.7 million estimate represents the full-time unemployed plus the full-time equivalent of those who were working involuntarily less than a full week. This estimate does not take into account the unemployed who are not counted in the government survey.

²¹ *Economic Report of the President* (Jan., 1969), p. 252.

²² U.S. Arms Control and Disarmament Agency, *Economic Impacts of Disarmament* (Washington, D.C., 1962).

welfare program would have its effect. But it is by no means obvious that these counteractions would have anywhere near the same impact on the economy as defense spending.

Economists are to a large measure captives of the neat models they create, and they consequently ignore strategic dynamic elements which keep the economy going. For example, they tend to underestimate, if not ignore, the special effects of persistent inflation on business practices regarding inventory accumulation and investment in plant and equipment. Even more important is the almost total neglect of the influence of stock market and real estate speculation on business investment decisions and the buoyancy of the especially important luxury trades. Inflation and speculation—partners of militarism—have been key triggers of our postwar prosperity, and they are too easily ignored as economists blandly transfer a block of GNP from one category to another, as if such transfers are made in the economy as simply as one keeps accounts.

The experience of the last depression still remains a challenge to economists to come up with an explanation of the way in which the economy operates in reality. For example, consider where we stood in 1939 after ten years of depression. Personal consumption expenditures had finally climbed to a new high—6 percent above 1929 in constant prices. Yet, at the same time, nonresidential fixed investment expenditures were 42 percent below the level of 1929, and residential construction was 20 percent below.²³ Despite six years of rising consumer spending and the start of orders flowing in from France and England for rearming, the investment community was still in a state of depression, and over 17 percent of the labor force was unemployed.

In this connection, it is important to recognize that one of the major attributes of the huge military spending in the postwar years is its concentration in the producers durable field and the stimulation it gives to the demand for machinery and equipment. If we combine the spending for producers durable goods resulting from the military with private spending for the same type of goods, we find the following: 36 percent of the

output of the producers durable goods industries is purchased directly or indirectly by the federal government.²⁴ (These data are for 1963, before the impact of the Vietnam war.) It is here, I suggest, that we find the unique role of military spending in raising and sustaining production and employment at new highs.

There are, to be sure, other impacts of defense spending that help to explain the magnitude and structure of the postwar economy: the unique role of research stimulated and financed by military and space programs; the special place of defense spending in nurturing the growth and prosperity of key giant industrial and financial enterprises; the support given by U.S. military power to acceptance of the U.S. dollar as an international currency;²⁵ the ease with which military orders can be fed into the economy in spurts which act as adrenalin injections to the private sector.

At the least, it can be concluded, economic theory and analysis which omit imperialism and militarism from their underlying paradigm are far removed from the reality of today's world. More realistically, it can be said that they operate to obscure the truth about the great problems and dangers of the second half of the twentieth century.

²⁴ Calculated from tables in "Input-Output Structure of the U.S. Economy: 1963," *Survey of Current Business*, Nov., 1969. The percent of direct and indirect output attributable to (a) gross private fixed capital formation and (b) federal government purchases were used. These percentages were applied to the gross output of each of the industries manufacturing durable goods. It is generally estimated that 85 percent of federal government purchases are for the military. The figure is probably higher for durable goods manufacturing industries alone.

²⁵ Given the inadequate U.S. gold reserve, the U.S. dollar can serve as an international currency only as long as foreign banks are willing to keep dollar credit balances in the United States as a substitute for gold payments. It is interesting that former Under Secretary of the Treasury for Monetary Affairs Robert Roosa included the military strength of the United States as a factor in maintaining the present international monetary system: "Moreover, the political stability and enormous economic and military strength of the United States have also increased the desirability of keeping balances here rather than in any other country in the world." (Robert V. Roosa, *Monetary Reform for the World Economy*, 1965, p. 9.)

²³ *Economic Report of the President* (Jan., 1969), p. 228.

DISCUSSION

STEPHEN HYMER: As Harry Magdoff reminds us, quoting Rosa Luxemburg, neoclassical economics deals with market relations and not with power relations. The study of imperialism, in contrast, is mainly concerned with the level of coordination above that of the market where state power is used to manipulate the economic framework within which supply and demand interplay. The analytical focus is the way one country exercises power over another and how this affects trade, development, and the distribution of income.

The first point to be stressed is that the neoclassical model which includes market equations and excludes political equations is misspecified (to use econometrics terminology) and yields biased estimates and wrong predictions. The comfortable assumption that one can concentrate on economic relations and leave the analysis of power to other disciplines is not tenable when one admits, as who could deny, the crucial role of the state in shaping the economy through its policies on infrastructure, education, production, etc. The cost of ignoring political factors in these circumstances is an inability to identify economic relations and therefore an inability to make policy recommendations. (Marxist economists, it should be noted, are no less free from this deficiency than neoclassical economists. Witness the studies trying to prove breakdown or exploitation by reference to capitalist market behavior—e.g., lack of demand, etc.—rather than by class struggle.)

This panel on imperialism is significant because it represents an attempt, albeit a limited one, to introduce power relationships into the analysis of international economic exchange, and thus to help reduce the biases inherent in the purely "economic" models of international trade usually presented at sessions of the American Economic Association.

It is more easy to recognize the importance of political factors the further one goes back in history. Every one would admit that international trade during the mercantilist period did not fit the decentralized market model by any stretch of the imagination. Rather, it was characterized by alliances of monarchs and merchants, using their superior power to murder and enslave large numbers of people in the hinterland in order to extract a surplus for the benefit of the few in the metropole.

In the second international economy—the age of European imperialism between 1870 and 1939—the market played a much more important role but nonetheless, state power continued to be decisive in regulating conditions of exchange between countries. The fact that England's industrial revolution was imitated by only a few other countries and that the gap between the industrialized and nonindustrialized countries maintained itself or widened despite the very great expansion of trade, depended crucially on government policy of the period and cannot be explained without reference to imperialist relations.

In those countries directly or indirectly dominated

by metropolitan powers, governments, instead of promoting development, pursued a policy of promoting cheap labor and channeling it towards exports (by removing employment alternatives, by biasing infrastructure, and by protecting metropolitan enterprise against local enterprise). This policy, reflected in the highly elastic supply of primary products at low prices, acted to extract a surplus from the population and to transfer it in part abroad and in part to ruling elites. Force played an important role in creating this structure and maintaining it. In short, many of the problems of underdevelopment—lack of infrastructure, lack of human capital, lack of domestic entrepreneurship, lack of a modern government administration, and lack of national political traditions—can be traced to imperial relationships, and that this aspect should be neglected in the current literature on economic development and international trade is inexplicable on scholarly grounds.

What about the modern period? The underdeveloped countries have achieved national independence and presumably can control their government expenditure in order to achieve more development at home. However, these papers have argued that this does not mean that imperialism—the control of the less developed countries by the advanced—is dead. They have rather tried to show that: (1) the interest of the developed countries in the markets, raw materials, and investment opportunities of the underdeveloped world has not abated; (2) the use of coercive military power by the dominant countries to control the underdeveloped countries is very large; and (3) economic trends point to an intensification, or at least a maintenance, of dependency relationships during recent years.

However, it is clear that form has changed and much work remains to be done if the content of the new imperialism is to be properly understood. The most important feature of the present international economy is that the leading business enterprises have outgrown the national economy and are global in outlook and interest, though not, of course, in ownership.

This trend towards multinational corporations breaks the simple connection between the nation state and the national bourgeois that characterized the old imperialism. In many cases, business enterprises (the microsm) are now broader than nation states (the macrocosm). In other words, the part is greater than the whole and old balances and rivalries have, to some extent, been upset. The corporation gains flexibility by crossing international boundaries and erodes the power of the nation state. But state power is crucial to the functioning of the economy. The question then arises as to how it will be mobilized and exercised in the future if the powers of nation states are weakened.

Will the corporations be able to construct supranational states commensurate with their own multinational span of control? Or, will the system flounder

because economic integration by business enterprise has proceeded much faster than political integration by states. For example, it is well known that the use of monetary and fiscal policy to control the economy becomes less effective the more open the economy and the greater the degree of foreign investment. Are we approaching a stage when each nation state is too weak to pursue an appropriate stabilization policy? If so, we may well be close to a serious economic crisis unless the Keynesian problem of aggregate demand has somehow been magically solved by the advent of the multinational corporation or unless political integration on a world level speeds up considerably.

To take another example, the present plans to create a new world division of labor allowing the underdeveloped countries an expanded role in the exports of manufactures requires a number of important policy measures in the advanced countries, ranging from an expanded aid program to finance education and urbanization in the hinterland, to a removal of restrictive practices to allow labor-intensive manufactures to enter the developed world. This would seem to require more political cohesiveness in the advanced world than now exists. Yet, if these measures are not undertaken, growth in the underdeveloped world may be too slow to maintain political stability without severely repressive regimes. The system may not be viable.

These are burning questions. Unfortunately the limited horizons of neoclassical economics and the limited application of Marxist economics puts scholars in a very weak position in trying to be of some help in answering them.

VICTOR PERLO: Mr. Magdoff discusses the leap forward in U.S. militarism since World War II as U.S. imperialism has taken on the objective of global domination.

I wish to emphasize the decisive character of that leap. In 1880, there were 38,000 men in the armed forces, or 0.07 percent of the population. In 1898, as the U.S. carried out its first imperialist war of aggression, there were 236,000 men in the armed forces, or 0.3 percent of the population. There was little change thirty years later. Today there are 3½ million men, or 1.7 percent of the population. Similarly, in 1880, the military budget amounted to 0.5 percent of the GNP, in fiscal 1899 to 1.5 percent, and today to about 9 percent. Until recent decades the bulk of government spending was by state and local governments, and government economic activity was generally slight. So 50 percent of the federal budget could be for military purposes and still have only a minimal economic effect.

Mr. Magdoff says U.S. militarism is used "to open and keep open the doors for trade and investment in other parts of the world." I would like to state that more sharply, and to emphasize a qualitative change in the objective. The special international economic characteristic of modern imperialism is the derivation of superprofits by giant corporations from the exploitation of labor of people abroad, principally in areas controlled or heavily influenced by the given imperialist power.

Lenin's point—that the export of capital becomes especially important in relation to the export of goods—has reached an unprecedented degree of development in the case of U.S. imperialism, so that goods produced in enterprises owned by U.S. corporations in other countries now exceed \$100 billion yearly, or three times the annual export of goods from the United States.

Superprofits from foreign investments are greater than in the early part of the century, when colonial investments consisted mainly of extractive and agricultural industries with relatively low technical levels. U.S. imperialism—to some extent even formerly, but especially in recent decades—emphasizes combining the most advanced techniques with the most ruthless imposition of precapitalist wage levels and oppression of labor. We thus have the great electronic corporations setting up factories in U.S.-occupied Taiwan and South Korea, at wages of \$15-\$20 per month and with labor productivity above that prevailing in similar establishments in the United States. The rate of exploitation of labor is of the order of 1,000 percent or more.

The connection between militarism and superexploitation is illustrated by *Fortune's* story about promoter Herbert Fuller, who wanted to put up a \$10 million sugar mill in South Vietnam. The people who live at the desired location do not want him. Says *Fortune*: "When the troops arrive to clear the area, as they sooner or later must, this American capitalist will literally be one step behind them. 'I am in it for the money,' Fuller says. 'We could get back our investment in two years.'"

Magdoff notes how World War II pulled the United States out of the Great Depression, that militarism has been a major economic and scientific-technical stimulator. Militarism has also figured in the domestic political-economic strategy of monopoly capitalism, to turn the country in a reactionary direction, and increase the rate of exploitation of labor at home.

Militarism and anti-communism have made it possible to shift the tax burden radically in the past several decades from capital to labor.

In this environment real wages per unit of output in industry have declined dramatically over the past decade; the contrasts between the super-rich and the impoverished tens of millions have become more glaring than ever. Savage domestic racism has become more profitable as many millions of black people, Spanish-speaking people, Indians, etc., have been forced into the urban ghettos.

Many radicals think that the superprofits of imperialism have been used to bribe and co-opt American labor into the great affluent middle class, and permanently blunt the capital-labor conflict. I think the costs to labor have greatly exceeded any possible trickle down benefits from foreign investment profits, that events today show and will show more clearly that the capital-labor class struggle remains an essential derivate of the social structure, ultimately aggravated by imperialist militarism.

Magdoff implies that ending the stimulus of militarism might plunge the country back into conditions

like the depressed 1930's. Here I must differ sharply. Briefly, a vital element of modern imperialism is the phenomenon designated by Lenin as state-monopoly capitalism—the interpenetration and extremely close collaboration of government and big business in all spheres of life. This has developed very far in economic affairs in recent decades. It has changed the character of economic crises and cycles and made most unlikely a repetition of 1933. Militarism is by no means the only economic regulator of state-monopoly capitalism, nor is it the most important regulator in most imperialist countries that are doing much better than the United States economically.

As the world has evolved, militarism has become a handicap to economic progress, a source of economic difficulties and financial crises.

I quote figures of the U.S. Arms Control and Disarmament Agency on the percentage of GNP going to military purposes in the four leading capitalist countries. The U.S. heads the list with 8.5 percent, Britain follows with 5.8 percent, West Germany with 4.1 percent, and Japan is way down with 1.0 percent. Now we consider the unemployment percentage in the same countries in the same order: The U.S. 3.6 percent, Britain 2.5 percent, West Germany 1.6 percent, Japan 1.2 percent.

In measures of economic welfare, such as the increase in real wages and of the rate of economic growth, we find a marked inverse correlation with the percentage of militarism. The harmful effects of extreme militarism on the U.S. economy have become so widely recognized that there has developed a significant congressional opposition to the military-industrial complex, and the stock market, contrary to some earlier tendencies, is now bullish on peace news and bearish on war news.

It has become clear that none of the extremely critical problems of internal life—declining real wages of most workers, deterioration of public facilities, inflation, record tight money, the fantastic deficit in international payments, pollution, racism—none can be eased until there is an end to the Vietnam war and a reduction in militarism.

Moreover, we are right now going into the first real wartime recession in U.S. history, with some establishment economists fearing that it may be serious and accompanied by a financial panic.

I agree with the view that none of these problems can be fully or permanently solved under capitalism, with or without militarism. But the fight against militarism is essential to ease the harmful effects of these problems on working people, to create more favorable conditions for radical solutions, and for the ultimate establishment of socialism in the United States.

Magdoff associated rampant militarism in the U.S. with the desire to overthrow or contain socialist nations and to thwart national liberation movements. But he fails to show the interconnection between these motives. Why did the U.S. require only 100,000 men and \$150 million to defeat Spain and suppress the Cuban, Puerto Rican, and Philippine people, while with 600,000 men and \$30 billion per year it cannot defeat the Vietnamese people?

The difference lies in the higher stage of national liberation struggles today, in the combination of the fight for political independence with that for basic social revolution, in the leadership of Communists, in the existence and international solidarity of socialist states. The massive military and economic aid of socialist countries, first the U.S.S.R., has been essential in providing the Vietnamese people the kinds of material with which they have been so successfully resisting U.S. imperialist aggression, just as a Communist orientation and Soviet economic and military assistance have been an essential ingredient in helping the Cuban people maintain their newly won independence from the furious and continuing counterassault of U.S. imperialism.

Nor is this vital aspect of the contemporary world inconsistent with the aim to combat militarism. The essential aim of diplomatic negotiations for disarmament, as of the peace movement's domestic campaign against the industrial-military complex, is to create conditions in which all peoples can determine their own destinies—including social structure, rules about foreign investment—without the waging of military and economic warfare against them by any great power. And to free the people of the great powers from the dangers and burdens of militarism.

A final point. Lenin described imperialism as the final stage of capitalism, as moribund, decaying capitalism, rotten ripe for the socialist revolution.

The extreme militarism of U.S. imperialism is a sign of weakness, not strength. Genocidal warfare, whether conducted by Hitler in Europe or by Johnson and Nixon in Asia—the waving of H-bombs—are the desperate last thrusts of dying empires. They are no less dangerous for that reason. But look at the scorecard of imperialism during the twenty-five years of U.S. leadership:

1. The number of countries professing socialism, having Communist or similar parties leading, and having public ownership of the main means of production has increased from one to fourteen, with over a billion people.

2. The area of the colonial world—the most secure sphere of imperialist domination—has shrunk to a fraction of its former size.

3. More and more developing countries are striving for noncapitalist development, are restricting or expelling foreign corporations from dominating positions, ending imperialist military and intelligence bases, relying more and more on trade and aid relationships with the socialist world, setting an ultimate socialist goal.

4. The trade-union movements, revolutionary parties, and peace movements have multiplied numerically in the major industrial countries. Working people have won numerous social and economic reforms. They can also win the most important reform of all—by turning back the war danger associated with imperialist militarism. This is essential towards accomplishing the historically inevitable transition from imperialism to socialism without a holocaust.

ARTHUR MACEWAN: An explanation of imperialism requires an analysis of the tendencies toward the geo-

graphic expansion of capitalist economies. Richard Wolff has provided a good review of the various bases for expansion, and Harry Magdoff has provided documentation of the role of the state in that process. I would like to extend their arguments by considering some aspects of power relations in an expanding capitalist system.

We need to explain first why expansion is so important that the state is willing to use its power—diplomatic power, “aid”-giving power, military power—and often incur high costs in order to assure a free rein for expansion. An explanation is not immediately obvious if one simply examines the size of trade and foreign investment relative to the aggregate operations of the economy. If one views the state as acting in a rational manner to further the general welfare of its nationals, without particular regard for their class, an economic explanation of the international activities of the state is not easily developed.

On the other hand, if the state is viewed as using its power to serve the interests of capitalists as a class, then an economic theory of imperialism begins to make sense. The impact of the international activities of U.S. business is far greater on profits than on aggregate demand. (Profits from foreign investments accounted for around 15 percent of total corporate profits in the U.S. during the mid-1960's.) Not only is foreign economic activity more important for those who earn their income from capital than for those who earn their income from labor, but the profits from foreign investment are highly concentrated in the largest firms. In 1966, for example, more than half of all profits remitted to U.S. corporations from abroad went to sixteen firms. All of these firms ranked among the top thirty in the *Fortune* listing.

This is not to say that the state's international activities can be interpreted as being designed simply to maximize aggregate profits of U.S. business (or of the largest firms). Nonetheless, these facts do begin to focus attention on the proper group; that is, on capitalists as a class.

Capitalists as a class are served well by the ordinary functioning of the capitalist system. Even without state intervention the institutions of the system assure that capitalists will be at the top in a society where income, power, and prestige are unequally distributed. While the most dramatic and often the most revealing operations of the state are its actions which directly and immediately serve the interests of business, the highest priority function of the state—both domestically and internationally—is the preservation and spreading of the institutions of capitalism. (In more popular jargon, the role of the state is to insure a “favorable investment climate.”)

This sort of theory of the state emerges directly from an analysis of the historic development of capitalism and, in particular, from an analysis of the state's relation to the process of capital expansion. In

the early period of capitalist development, the primary problem was that of creating a nation or, more to the point, of breaking down restrictions on the labor market and on capitalist organization of the work process. At a later stage in the process of spreading the system, the state took on the role of developing the “sphere of influence.” Today, when one capitalist nation has become dominant, the state uses its power to integrate the international capitalist system further.

This integration means at a minimum that the nefarious aspects of capitalism—inequality, alienation, destruction of the environment—are spread (or, insofar as they already exist, they are maintained). However, because this integration takes place under the dominance of the business interests of an advanced capitalist nation, the output-expanding capacity of capitalism is not necessarily transmitted to the poor countries. It is this dominance and its consequences which are the concern of Theotonio Dos Santos. The dominance itself is assured by the power of the imperialist state, but the dominance has its everyday impact through the economic sphere. Dos Santos has cited, for example, the price relations of the international market and the monopoly power of business from the imperialist nation and how they affect underdevelopment.

A further aspect of the “economic” dominance is the effect which imperialism has upon class and power relations in the poor countries. Because of its power, which operates both within and outside the market, U.S. business (or business from other advanced capitalist nations) is able to preempt investment opportunities and inhibit the development of an historically progressive industrial bourgeoisie in the poor countries. For example, the “balance-of-payments crises” endemic to poor countries tend to work in the favor of foreign as opposed to domestic capital. The former always has a source of foreign exchange. Also, foreign capital often enters the poor countries before there is any appreciable domestic industrial class to compete with, and alliances are formed, both political and economic, between the foreign interests and domestic groups which are antagonistic to industrialization; e.g., agrarian and military elites, commercial interests tied to foreign trade. Thus, foreign business helps reactionary groups in their control of the state and assures that policies favorable to a domestic industrial class will be inhibited. Under such circumstances, while some industrialization can take place, it will tend to be dominated by foreign interests and not lead to the development of the nation.

Thus, within the economy of the poor country, as well as within the international system, imperialism has its influence through class power relationships which operate both through the state and in the economic sphere. It would seem to me that analyses of imperialism should give explicit attention to those relationships.

DISTRIBUTION ISSUES: TRENDS AND POLICIES

POSTWAR CHANGES IN THE SIZE DISTRIBUTION OF INCOME IN THE U.S.*

By EDWARD C. BUDD
Pennsylvania State University

My assignment for this session is to review recent changes in size distribution in the U.S. I have interpreted "recent" to include the entire postwar period rather than just the past decade. Most analysts have concluded that, in contrast to the changes that appear to have occurred from 1929 to World War II, the relative distribution of income since the war has been quite stable.¹ While perhaps the only inference that the quality of our income size data will support, it may be worthwhile to take another look at the evidence we have.

The widely used Lorenz curve technique for representing relative size distributions may possibly contribute to the acceptance of the stability hypothesis. It is, after all, rather insensitive to all but major changes in distribution; we are all too prone to conclude that if the Lorenz curves for different years lie close together or if the income shares of various quantile groups change only by a percentage point or two, nothing of consequence has happened to inequality. While the Gini concentration ratio is perhaps the most useful—and certainly the most widely used—measure of changes in inequality, it does, as we know, produce an ambiguous measure of changes in inequality if the two relevant Lorenz curves intersect, and an intersection is implied if the rich and the poor lose relative to the middle of the distribution. Even if the curves do not intersect, one part of the distribution, e.g., any given quantile such as the 45th percentile or the 3rd decile, may gain or lose more, relative to their position in some earlier year, than does some other quantile.

A convenient way of revealing these changes is to compare the mean income of any quantile relative to the mean income of the distribution as a whole; this mean income ratio is simply the quantile's share of total income divided by the

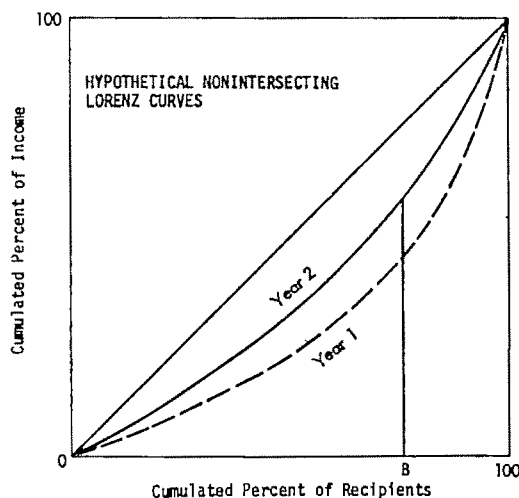


FIGURE 1

size of the quantile. In graphical terms, it is the slope of the chord connecting the lower and upper points on the Lorenz curve for the quantile, or, if the quantile is sufficiently small, the slope of the Lorenz curve itself. In the ensuing discussion this ratio will be referred to as the "relative mean income."

The hypothetical behavior of relative mean incomes for different pairs of Lorenz curves is illustrated in Figures 1 and 2. In Figure 1, inequality is assumed to be less in year 2 for every part of the distribution, so that year 2's curve lies everywhere above year 1's. At the *B*th percentile, the relative mean incomes are equal; below *B*, year 2's exceeds year 1's; above *B*, year 2's is less than year 1's. Further, the ratio of year 2's mean to year 1's mean falls as we move from the lowest to the highest groups; at *B*, the ratio is one, and beyond that point, less than one. This behavior is of course consistent with our intuitive notion of a reduction in inequality: the incomes of lower-income groups rise, and those of upper-income groups fall, relative to the mean income of the entire distribution. In Figure 2, the Lorenz curves are assumed to intersect, indicating a re-

* I am indebted to Thomas C. Whiteman for assistance and to Mark Prichard for developing and writing the computer program. Inquiries concerning the computer program developed for this project may be addressed to the author.

¹ For example, [10, pp. 20-26]. For a contrary view, however, see [15].

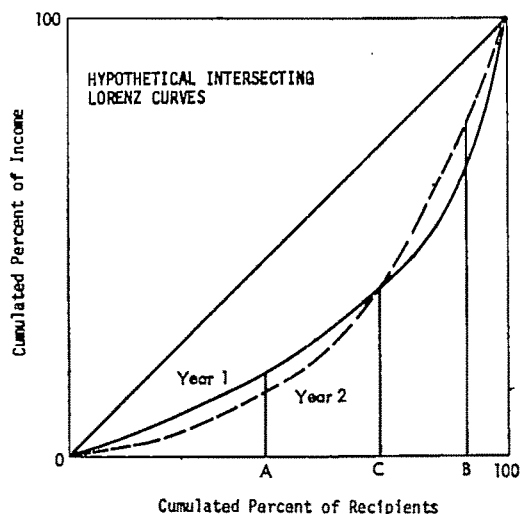


FIGURE 2

duction in inequality at the upper part of the distribution and an increase in inequality for the lower. Here there are two percentiles, the *A*th and the *B*th, at which the relative mean incomes are the same. From zero to *A*, year 2's mean is less than 1's; the ratio of the mean of year 2 to that of year 1 rises as we move from zero to *A*, equalling unity at *A*. Beyond that point, year 2's mean is greater than 1's, the ratio between them first increasing, reaching a maximum, then decreasing, until it equals unity at *B*. Beyond *B*, year 2's mean again exceeds year 1's.

It is, perhaps, easier to see these relationships in a graph due to Robert Schutz [14], which shows the relative mean income for each year's distribution as a function of the cumulated percent of recipients; the horizontal axis of this graph is thus the same as that of the Lorenz curve. Figure 3 illustrates the relative mean income curves underlying the Lorenz curves shown in Figure 2. If the existence of loss incomes were assumed, the Lorenz curves would intersect the horizontal axis, and the relative mean income curve would assume negative values to the left of the point at which the slope of its corresponding Lorenz curve were zero.

A useful way of introducing the problem of changes in inequality in the postwar period is by way of examining the changes between 1946-47 and 1960-61 in the income shares of the quantiles used by the Office of Business Economics in its old income size series for consumer units (families and unattached individuals): the five quintiles, and the top 5 and 1 percent of recipi-

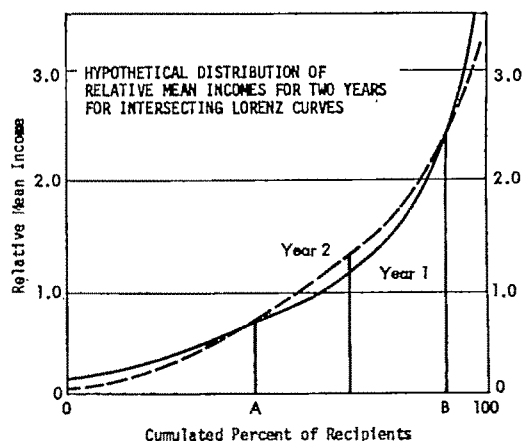


FIGURE 3

ents. The data are summarized in Table 1: the upper panel shows the quantile income shares for the averages of the above years, as well as for 1950-54 and 1955-59; the lower panel, the implied relative mean incomes. The changes in shares may not seem large, that of the bottom quintiles's, for example, falling by only .4 percentage points and that of the top 5 percent, by 1.5 percentage points. But they do imply changes in relative mean incomes that we might not deem insignificant; the relative mean income of the bottom quintile fell by 8 percent, of the next quintile, by 2 percent. The quantile encompassing the 40th to the 95th percentiles, on the other hand, experienced a rise of over 3 percent in its relative mean, while that of the top 5 percent fell by 7 percent, and the top 1 percent, by 14 percent.

Somewhat the same pattern is shown by the Survey Research Center's Survey of Consumer Finances [16]. While the rounding of the income share estimates for quantiles to whole numbers in the Center's publications prevents drawing any definite conclusions about changes in the relative mean income of the bottom two quintiles, it appears that, for the same set of years as used in the OBE comparison, the relative mean income of recipients encompassing the 40th to the 90th percentiles rose between 6 and 7 percent, while that of the top decile fell by about 11 percent. These percentage changes are not altered if the years 1960-61 are replaced by more recent years, such as 1965-66.

On the other hand, these changes are somewhat less evident in the quintile distributions for families and unrelated individuals combined, from the Current Population Survey. The income

TABLE 1
DISTRIBUTION OF FAMILY PERSONAL INCOME, SELECTED YEARS
OFFICE OF BUSINESS ECONOMICS
QUANTILES

	Quintiles				Next 15%	Top 5%	Top 1%
	Lowest	2nd	Middle	4th			
A. Income share (percent)							
Average of							
1946-47	5.0	11.1	16.0	21.9	24.9	21.1	8.6
1950-54	4.9	11.2	16.4	22.4	24.6	20.5	8.4
1955-59	4.7	11.1	16.3	22.4	25.3	20.1	7.7
1960-61	4.6	10.9	16.3	22.7	25.9	19.6	7.4
B. Relative mean income							
Average of							
1946-47	.25	.555	.80	1.095	1.66	4.22	8.60
1950-54	.245	.56	.82	1.12	1.64	4.10	8.40
1955-59	.235	.555	.815	1.12	1.69	4.02	7.70
1960-61	.23	.545	.815	1.135	1.73	3.92	7.40
Percent change, 1946-47 to 1960-61	-8	-2	+2	+4	+4	-7	-14

SOURCE: [11].

shares show some year-to-year variability, and the conclusion drawn depends to some extent on the years selected for comparison. If we take the years 1947-48 and 1960-61, a similar Lorenz curve intersection phenomenon is apparent.² The relative mean income of the two bottom quintiles fell by over 5 percent, that of the next group comprising the 40th to the 95th percentiles rose by 2 percent, and that of the top 5 percent fell by 2 percent.

Since the evidence seems to point to intersection of Lorenz curves in postwar years, it was decided to investigate the matter further by constructing a smooth Lorenz curve for each published distribution, so that relative mean incomes could be determined for every percentile of the distribution and the distributions could be inspected for intersection of their corresponding Lorenz curves. Published size distributions give frequencies for dollar income size brackets that remain relatively constant from year to year; as a result, sizes and positions of the quantile readings for relative distributions derived from them

² Data are from [10, p. 21]. Income shares for the combined distribution of families and unrelated individuals have not been published for years following 1962. The year 1948 was substituted for 1946 (the year used in the OBE comparison), since there was no CPS for the entire population in the latter year.

vary considerably. In order to express the relative distributions in terms of uniform quantiles, e.g., quintiles, resort must be had to interpolation formulas or the freehand drawing of Lorenz curves.

Before Lorenz curves can be constructed, it is necessary to have an estimate of the mean income or the income aggregate as well as frequencies within each class interval. The CPS is virtually alone in failing to provide this information, at least for years prior to 1967. One of the difficulties in assuming that class interval means for all but the top part of the distribution fall at the midpoint is that respondents tend to report their incomes in rounded numbers, particularly in multiples of \$1,000 or \$500. Since these "round-off" points tend to correspond with the lower limit of a Census class interval, the mean is normally less than the midpoint. Instead of assuming, therefore, that the mean corresponds with the midpoint for all intervals below the open-end one, with a Pareto function used to estimate the mean of the latter,³ I decided to use, separately for

³ Census uses the midpoint for its computations of aggregates [6, pp. 33-34]; Schultz used the geometric mean of the endpoints of the interval [13, pp. 76-77]. These references also describe the use of the Pareto function for estimating the mean of the open-end interval, and the means assumed for that interval if the function cannot be applied in any given year.

families and unrelated individuals, the actual class interval means for the 1964 CPS that we have computed in our work at OBE. A test was made to see if these means were altered by assuming that aggregate income was three-fourths, and then one-half, of the 1964 aggregate, in order to take account of the fact that incomes reported in the CPS have increased through the years. The results of this test were essentially negative. The bottom open-end interval also presents a problem because of the inclusion of no income and negative income cases; in order to compute the mean for the bottom interval for years prior to 1964, the proportion of these frequencies to the total number of recipients was held constant at the 1964 proportion.⁴

A variant of what might be considered an interpolation or smoothing device was used for this study. Polynomial functions were fitted to the observations ranging from the bottom up to the 70th, 75th, 80th, and 85th percentiles, respectively, with the final choice being for a fourth degree polynomial encompassing observed points up to the 80th percentile. Only occasionally did the predicted value of income differ from the observed value by as much as a tenth of a percentage point.

The polynomial could not be used, however, for the top of the distribution. This should not be surprising: we know that it is virtually impossible to describe empirical distributions accurately by just one function. The Pareto function, for example, fits well only the top 20 percent of recipients or less; the lognormal function, judging by my experience in my OBE work, does not accurately predict frequencies for either the top 5 or 10 percent of recipients or the bottom quintile, whether distributions of total income or of individual income types are used. For the top observations, resort was therefore had to a function originally suggested by Corrado Gini many years ago: $Y_i = B_i X_i^{2.5}$; where Y_i is the i th percent of recipients cumulated from the top, and X_i , the percent of total income they receive [8, p. 75]. The fit finally selected was for those readings encompassing no more than 30 percent of the top recipients, although the function was also fitted for one additional observation above that limit.

⁴An inspection of those tax returns reporting Schedule C, Schedule F, partnership, or rental income indicated that the ratio of those returns reporting a loss income to those reporting a positive amount of these income types did not decline during the postwar period; if anything, it rose somewhat. If the latter were the case, the procedure used would result in a small understatement of the mean of the bottom interval for earlier years.

The two equations—the polynomial and the exponential—were then solved for their intersection point, which usually occurred between the 74th and 80th percentile of recipients. For most years and distributions the exponential cut the polynomial from below, implying a steeper slope for the former at the point of intersection. This difference in slopes implies a small discontinuity in the relative mean income function at the intersection point, although the difference was never more than a tenth of one point, and normally less than that. In the table giving relative mean incomes for various percentiles (Table 3), it should be noted that the relative mean might have differed from itself by as much as 9 percent, e.g., 1.20 rather than 1.30, had the intersection point occurred just above rather than just below that percentile. As a check for the reader, the point of intersection, and the difference in the relative mean incomes for the two functions at that point, in terms of recipient units, are given in the last two columns of Table 6. The computer then read off the cumulated share of income for each percentile of each distribution and the relative mean income for that percentile. I might add that the selection of those functions for each year that would give a good fit to the data and could also be spliced together with least damage to the relative mean income function could have been somewhat improved had it not been necessary, in order to economize on computer and programming time, to write one general rule for the computer to follow for selecting and combining the relative functions for each year.

The computer results for each of the distributions are summarized in Tables 2 and 3 for each fifth percentile, corresponding to the 10th, 15th, etc., and for the 99th. The 5th percentile was omitted because of misgivings, not only about the quality of the fit at the very bottom, but the quality of the data on which it is based as well. Perhaps the results should not be taken too seriously below the 15th or 20th percentiles. Table 2 shows the cumulated percent of total income for each of the percentiles selected, and Table 3, the relative mean income for that percentile.

The only "check" that can be made of the values in Table 2 is to compare them with quantile estimates from published sources, which are also known to be interpolated—although we admittedly do not know whose interpolation method is, in some sense, better. I did compare my cumulative share estimates with those made by OBE for quintiles and the top 5 percent of its distributions for consumer units for the years 1947, 1950, 1955, 1960, and 1961. Only in 1950

TABLE 2
CUMULATED SHARE OF INCOME, AT SPECIFIED PERCENTILES

	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	99%
CPS: FAMILIES + UNRELATED INDIVIDUALS: MONEY INCOME BEFORE TAX																			
1944 CPS	0.8	1.9	3.3	5.2	7.5	10.2	13.3	16.8	20.7	25.2	30.0	35.3	41.2	47.5	54.4	62.0	70.6	81.1	93.2
1945 CPS	0.8	2.1	3.8	5.9	8.5	11.5	14.8	18.5	22.6	27.1	32.0	37.3	43.0	49.2	56.0	63.6	72.2	82.4	93.9
1947 CPS	0.7	1.8	3.4	5.4	7.8	10.7	13.9	17.5	21.5	25.9	30.7	35.9	41.6	47.7	54.5	62.1	70.7	81.1	93.2
1948 CPS	0.7	1.8	3.4	5.4	7.9	10.8	14.1	17.8	21.8	26.3	31.2	36.5	42.2	48.3	55.1	62.6	71.1	81.4	93.3
1949 CPS	0.6	1.7	3.2	5.2	7.6	10.4	13.7	17.3	21.4	25.9	30.8	36.1	41.9	48.1	55.0	62.7	71.3	81.7	93.5
1950 CPS	0.6	1.6	3.0	5.0	7.4	10.2	13.5	17.2	21.3	25.8	30.8	36.1	41.8	47.9	54.8	62.5	71.1	81.5	93.4
1951 CPS	0.6	1.7	3.3	5.4	7.9	10.9	14.3	18.1	22.3	26.9	31.9	37.3	43.0	49.3	56.1	63.6	72.1	82.2	93.8
1952 CPS	0.6	1.8	3.4	5.5	8.0	11.0	14.3	18.1	22.2	26.7	31.7	37.0	42.9	49.3	56.0	63.6	72.1	82.2	93.8
1954 CPS	0.5	1.5	3.0	4.9	7.2	10.0	13.2	16.9	21.0	25.6	30.6	36.1	42.0	48.3	55.3	63.0	71.8	82.2	93.9
1955 CPS	0.6	1.7	3.2	5.1	7.6	10.4	13.7	17.5	21.7	26.3	31.3	36.8	42.7	49.0	56.0	63.7	72.3	82.6	94.0
1956 CPS	0.6	1.7	3.3	5.3	7.8	10.7	14.0	17.8	22.0	26.6	31.7	37.1	43.1	49.4	56.4	64.1	72.6	82.8	94.2
1957 CPS	0.7	1.6	3.0	4.9	7.2	10.0	13.3	17.1	21.4	26.2	31.4	37.1	43.1	49.6	56.6	64.4	73.0	83.2	94.4
1958 CPS	0.7	1.7	3.2	5.2	7.7	10.6	13.9	17.6	21.8	26.5	31.5	37.0	42.9	49.3	56.3	64.2	72.9	83.2	94.4
1959 CPS	0.6	1.7	3.2	5.1	7.5	10.4	13.6	17.3	21.5	26.1	31.1	36.6	42.5	48.9	55.9	63.6	72.3	82.6	94.1
1960 CPS	0.6	1.7	3.2	5.1	7.5	10.4	13.6	17.3	21.5	26.0	31.0	36.5	42.4	48.7	55.8	63.4	72.1	82.3	93.9
1961 CPS	0.7	1.7	3.1	5.0	7.4	10.1	13.3	16.9	20.9	25.4	30.3	35.7	41.6	47.6	54.9	62.6	71.3	81.8	93.7
1962 CPS	0.7	1.8	3.3	5.3	7.7	10.5	13.8	17.4	21.5	26.0	31.0	36.4	42.4	48.9	55.9	63.7	72.5	82.8	94.2
1963 CPS	0.8	1.8	3.3	5.3	7.7	10.5	13.8	17.5	21.6	26.2	31.2	36.7	42.7	49.1	56.2	64.0	72.7	82.9	94.3
1964 CPS	0.8	1.8	3.4	5.3	7.7	10.5	13.8	17.4	21.6	26.1	31.1	36.6	42.5	48.9	56.0	63.8	72.6	82.9	94.3
1965 CPS	0.9	1.9	3.5	5.5	7.9	10.7	13.9	17.6	21.7	26.3	31.3	36.8	42.7	49.1	56.2	63.9	72.5	82.8	94.2
1966 CPS	0.9	2.1	3.6	5.6	8.0	10.9	14.2	17.9	22.0	26.6	31.6	37.1	43.0	49.3	56.4	64.1	72.7	82.9	94.2
1967 CPS	0.9	2.1	3.7	5.7	8.1	10.9	14.2	17.9	21.9	26.5	31.4	36.8	42.7	49.1	56.1	63.7	72.3	82.5	94.0
1968 CPS	1.0	2.2	3.9	5.9	8.4	11.3	14.7	18.4	22.6	27.1	32.2	37.6	43.5	50.0	56.9	64.4	72.9	82.9	94.2
OBE: FAMILIES + UNATTACHED INDIVIDUALS: FAMILY PERSONAL INCOME																			
1947 OBE	1.5	3.0	4.9	7.2	9.9	12.8	16.1	19.6	23.4	27.6	32.0	36.9	42.1	47.7	54.0	60.9	69.0	79.1	91.6
1950 OBE	1.4	2.8	4.7	6.9	9.5	12.4	15.6	19.1	22.9	27.1	31.6	36.4	41.7	47.4	53.5	60.4	68.4	78.5	91.2
1955 OBE	1.4	2.9	4.9	7.2	9.9	12.9	16.3	19.5	23.9	28.1	32.7	37.7	43.0	48.8	55.1	62.0	69.9	79.8	92.0
1960 OBE	1.3	2.7	4.5	6.8	9.3	12.3	15.5	19.1	23.0	27.2	31.8	36.8	42.3	48.2	54.6	61.8	70.1	80.3	92.5
1961 OBE	1.3	2.7	4.6	6.8	9.3	12.2	15.4	19.0	22.9	27.1	31.7	36.6	42.1	48.0	54.4	61.6	69.5	80.2	92.5
BLS: SPENDING UNITS: MONEY INCOME AFTER TAX																			
1950 BLS	2.1	4.1	6.6	9.5	12.7	16.3	20.1	24.2	28.6	33.2	38.1	43.3	48.8	54.7	61.1	67.9	75.5	84.6	94.8
1960 BLS	2.1	3.9	6.2	8.9	12.0	15.4	19.1	23.1	27.5	32.2	37.1	42.5	48.2	54.4	61.1	68.1	75.9	85.1	95.1
IRS: ALL TAX RETURNS (ALL) & JOINT RETURNS (JNT): ADJUSTED GROSS INCOME																			
1948 ALL	1.3	2.6	4.4	6.6	9.1	12.0	15.2	18.8	22.6	26.8	31.3	36.2	41.4	47.1	53.2	59.6	67.2	76.9	89.8
1950 ALL	1.2	2.6	4.5	6.7	9.3	12.2	15.5	19.1	22.9	27.1	31.6	36.5	41.8	47.6	53.5	60.2	68.0	77.9	90.7
1952 ALL	1.0	2.2	3.8	5.8	8.2	10.9	14.1	17.6	21.5	25.8	30.5	35.6	41.2	47.3	53.6	60.8	69.0	79.3	91.9
1960 ALL	0.9	1.9	3.4	5.2	7.4	10.1	13.2	16.7	20.6	24.9	29.7	34.9	40.7	46.8	53.5	61.0	69.4	79.9	92.4
1964 ALL	0.7	1.6	2.8	4.5	6.6	9.1	12.0	15.4	19.2	23.5	28.2	33.4	39.3	45.6	52.3	59.9	68.5	79.2	92.0
1966 ALL	0.7	1.5	2.7	4.3	6.3	8.7	11.5	14.8	18.7	23.0	27.8	33.1	39.0	45.5	52.4	60.1	68.9	79.7	92.4
1967 ALL	0.8	1.6	2.8	4.6	6.3	8.6	11.5	14.8	18.5	22.8	27.6	32.9	38.7	45.0	51.9	59.5	68.3	79.1	92.0
1955 JNT	2.0	3.8	6.0	8.6	11.6	14.9	18.4	22.2	26.3	30.7	35.3	40.3	45.6	51.2	57.1	63.6	71.2	80.7	92.4
1960 JNT	1.8	3.5	5.6	8.2	11.1	14.4	17.9	21.8	25.9	30.3	35.1	40.1	45.5	51.3	57.5	64.2	71.9	81.4	92.9
1964 JNT	1.8	3.6	5.8	8.4	11.3	14.6	18.2	22.1	26.2	30.6	35.4	40.4	45.8	51.6	57.8	64.5	72.2	81.7	93.1
1966 JNT	2.2	4.0	6.3	9.0	11.9	15.2	18.7	22.5	26.6	30.9	35.5	40.5	45.8	51.6	57.8	64.7	72.1	81.2	92.5
1967 JNT	1.9	3.6	5.9	8.5	11.4	14.7	18.3	22.1	26.2	30.6	35.3	40.3	45.6	51.4	57.5	64.4	71.8	80.8	92.2

were the differences for cumulated income shares greater than .2 percentage points; in that year the maximum was .3 to .4 percentage points.⁵

From these two tables, the Lorenz curves for any two distributions can easily be compared. As an example, I have used the distributions of all tax returns (IRS) for 1948 and 1967, not because a comparison of these two are the most empirically significant, but because the distributions are sufficiently dissimilar to illustrate the approach. At approximately the median (the 51st percentile, to be exact), the slopes of the two distributions are both .81 and the cumulative difference in income shares reaches a maximum of 4.1 percentage points, with 1967 lying below 1948. In the 86th percentile, the two distributions cross, with a difference in their relative mean incomes of .25,

or 18 percent of the 1948 mean of 1.40. In the 98th percentile, the relative mean incomes are again equal (at a value of approximately 3.3) and the maximum cumulative difference in income at this percentile is equal to 2.2 percentage points, with 1948 exceeding 1967. The group comprising the 51st through (part of) the 98th percentiles thus gained at the expense of the bottom half and slightly more than the top 1 percent; the relative mean income of the former increased by 10 percent; that of the latter two groups fell by 18 and 20 percent, respectively.

Tables 4 and 5 are derived from Table 2. The first shows the income share of each of the groups—five percentiles wide—which lie between the chosen percentiles, with the exception that the first group comprises the bottom decile, and the last group, the top 1 percent. The second gives the relative mean income of the group; it is simply the share estimate in Table 4 divided by the size of the group as measured in percentiles.

A few words of explanation are in order on the final table from the computer printout (Table 6).

⁵ I have not included a comparison between my estimates for the CPS and the quintile estimates published by Census, since they would reflect not simply differences in the smoothing or interpolation method, but the difference as well in assumptions made concerning class interval means discussed previously.

TABLE 3
RELATIVE MEAN INCOME AT SPECIFIED PERCENTILES

	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	99%
CPS: FAMILIES - UNRELATED INDIVIDUALS: MONEY INC. BEFORE TAX																			
1944 CPS	0.16	0.24	0.32	0.41	0.49	0.57	0.65	0.74	0.83	0.92	1.01	1.10	1.20	1.31	1.44	1.59	1.83	2.32	3.76
1945 CPS	0.19	0.28	0.38	0.46	0.55	0.62	0.70	0.77	0.85	0.93	1.01	1.09	1.18	1.28	1.44	1.58	1.81	2.25	3.53
1947 CPS	0.16	0.26	0.35	0.44	0.52	0.60	0.68	0.75	0.83	0.91	0.99	1.08	1.17	1.27	1.43	1.58	1.82	2.31	3.76
1948 CPS	0.16	0.26	0.35	0.44	0.53	0.61	0.69	0.77	0.85	0.93	1.01	1.09	1.17	1.30	1.41	1.56	1.80	2.28	3.71
1949 CPS	0.15	0.23	0.34	0.43	0.52	0.60	0.68	0.77	0.85	0.93	1.01	1.10	1.18	1.27	1.45	1.59	1.83	2.30	3.66
1950 CPS	0.14	0.24	0.33	0.42	0.51	0.60	0.69	0.78	0.86	0.94	1.02	1.10	1.17	1.25	1.44	1.59	1.83	2.31	3.69
1951 CPS	0.16	0.26	0.36	0.45	0.54	0.63	0.71	0.80	0.87	0.95	1.03	1.10	1.18	1.31	1.42	1.56	1.79	2.24	3.56
1952 CPS	0.16	0.27	0.36	0.45	0.54	0.62	0.70	0.78	0.86	0.94	1.02	1.11	1.20	1.30	1.42	1.57	1.80	2.25	3.55
1954 CPS	0.14	0.23	0.32	0.42	0.51	0.60	0.68	0.77	0.86	0.95	1.04	1.13	1.22	1.31	1.47	1.62	1.85	2.29	3.58
1955 CPS	0.15	0.25	0.34	0.43	0.52	0.61	0.70	0.78	0.87	0.96	1.04	1.13	1.21	1.30	1.46	1.60	1.82	2.26	3.50
1956 CPS	0.15	0.25	0.35	0.44	0.53	0.62	0.70	0.79	0.87	0.96	1.04	1.13	1.22	1.31	1.45	1.59	1.81	2.24	3.46
1957 CPS	0.14	0.22	0.31	0.41	0.51	0.60	0.70	0.80	0.90	0.99	1.08	1.16	1.24	1.31	1.47	1.61	1.82	2.23	3.39
1958 CPS	0.15	0.25	0.34	0.43	0.52	0.61	0.70	0.79	0.87	0.96	1.05	1.13	1.22	1.31	1.49	1.63	1.84	2.25	3.40
1959 CPS	0.15	0.24	0.34	0.43	0.52	0.61	0.69	0.78	0.86	0.95	1.04	1.13	1.23	1.32	1.46	1.61	1.83	2.26	3.50
1960 CPS	0.15	0.24	0.34	0.43	0.51	0.60	0.69	0.78	0.86	0.95	1.04	1.12	1.21	1.35	1.45	1.60	1.82	2.27	3.54
1961 CPS	0.15	0.24	0.33	0.41	0.50	0.58	0.67	0.76	0.84	0.93	1.02	1.11	1.20	1.30	1.47	1.62	1.85	2.31	3.64
1962 CPS	0.16	0.25	0.34	0.43	0.51	0.60	0.68	0.77	0.85	0.94	1.03	1.13	1.23	1.34	1.46	1.62	1.84	2.27	3.47
1963 CPS	0.16	0.25	0.34	0.43	0.51	0.60	0.69	0.78	0.86	0.95	1.04	1.13	1.23	1.32	1.47	1.61	1.83	2.25	3.44
1964 CPS	0.16	0.25	0.34	0.43	0.51	0.60	0.68	0.77	0.86	0.95	1.04	1.13	1.22	1.32	1.48	1.62	1.84	2.26	3.45
1965 CPS	0.17	0.26	0.35	0.43	0.51	0.60	0.69	0.77	0.86	0.95	1.04	1.13	1.22	1.32	1.47	1.61	1.83	2.25	3.46
1966 CPS	0.17	0.26	0.35	0.43	0.52	0.60	0.69	0.78	0.86	0.95	1.04	1.13	1.22	1.31	1.46	1.60	1.82	2.24	3.44
1967 CPS	0.17	0.26	0.35	0.43	0.52	0.60	0.68	0.77	0.85	0.94	1.03	1.12	1.21	1.31	1.45	1.59	1.81	2.25	3.51
1958 CPS	0.19	0.28	0.36	0.45	0.53	0.62	0.70	0.78	0.86	0.95	1.03	1.12	1.25	1.33	1.43	1.57	1.78	2.21	3.43
OBE: FAMILIES - UNATTACHED INDIVIDUALS: FAMILY PERSONAL INCOME																			
1947 OBE	0.24	0.34	0.42	0.49	0.55	0.61	0.67	0.73	0.79	0.85	0.92	0.99	1.08	1.17	1.29	1.46	1.73	2.29	4.05
1950 OBE	0.23	0.32	0.40	0.47	0.54	0.60	0.67	0.73	0.79	0.86	0.93	1.00	1.09	1.16	1.27	1.44	1.72	2.29	4.14
1955 OBE	0.24	0.34	0.42	0.50	0.57	0.63	0.69	0.75	0.82	0.88	0.95	1.02	1.10	1.20	1.30	1.44	1.69	2.23	3.93
1960 OBE	0.22	0.32	0.40	0.48	0.55	0.61	0.68	0.74	0.81	0.88	0.95	1.03	1.12	1.22	1.34	1.51	1.77	2.29	3.87
1961 OBE	0.22	0.31	0.40	0.47	0.54	0.61	0.67	0.74	0.80	0.87	0.95	1.03	1.12	1.22	1.33	1.52	1.78	2.30	3.89
BLS: SPENDING UNITS: MONEY INCOME AFTER TAX																			
1950 BLS	0.33	0.44	0.53	0.61	0.68	0.74	0.79	0.84	0.89	0.94	1.00	1.06	1.13	1.22	1.29	1.42	1.61	2.06	3.16
1960 BLS	0.31	0.41	0.49	0.57	0.64	0.71	0.77	0.83	0.89	0.96	1.03	1.10	1.18	1.27	1.37	1.45	1.64	2.00	3.01
IRS: ALL TAX RETURNS (ALL) & JOINT RETURNS (JNT): ADJUSTED GROSS INCOME																			
1948 ALL	0.22	0.31	0.39	0.47	0.54	0.60	0.67	0.73	0.80	0.86	0.93	1.00	1.08	1.16	1.25	1.35	1.63	2.25	4.30
1950 ALL	0.22	0.32	0.40	0.48	0.55	0.61	0.67	0.74	0.80	0.86	0.93	1.01	1.09	1.19	1.23	1.40	1.68	2.27	4.20
1955 ALL	0.18	0.27	0.35	0.43	0.51	0.58	0.66	0.73	0.81	0.89	0.97	1.06	1.15	1.22	1.34	1.50	1.77	2.32	4.03
1960 ALL	0.16	0.24	0.32	0.40	0.48	0.56	0.65	0.73	0.82	0.90	0.99	1.08	1.19	1.27	1.39	1.55	1.81	2.34	3.96
1964 ALL	0.13	0.21	0.29	0.37	0.45	0.53	0.62	0.71	0.80	0.89	0.99	1.09	1.20	1.29	1.41	1.58	1.84	2.40	4.09
1966 ALL	0.13	0.19	0.27	0.35	0.43	0.52	0.61	0.70	0.80	0.90	1.01	1.11	1.22	1.33	1.44	1.61	1.87	2.40	4.01
1967 ALL	0.13	0.19	0.26	0.34	0.42	0.51	0.60	0.70	0.79	0.89	1.00	1.10	1.20	1.31	1.43	1.60	1.87	2.42	4.11
1955 JNT	0.30	0.40	0.48	0.55	0.62	0.68	0.73	0.79	0.84	0.89	0.95	1.02	1.09	1.12	1.22	1.38	1.63	2.14	3.75
1960 JNT	0.28	0.38	0.46	0.54	0.61	0.67	0.73	0.79	0.85	0.91	0.97	1.04	1.11	1.19	1.25	1.40	1.64	2.14	3.65
1964 JNT	0.29	0.39	0.47	0.55	0.62	0.68	0.74	0.80	0.85	0.91	0.97	1.04	1.11	1.19	1.26	1.41	1.64	2.13	3.60
1966 JNT	0.32	0.41	0.49	0.56	0.62	0.68	0.74	0.80	0.85	0.91	0.97	1.04	1.11	1.19	1.26	1.41	1.64	2.13	3.60
1967 JNT	0.29	0.39	0.48	0.55	0.62	0.68	0.74	0.79	0.84	0.90	0.96	1.02	1.10	1.19	1.30	1.42	1.55	2.05	3.63

Gini himself used the value of the exponent B_2 in the exponential function as a direct measure of inequality. While this measure applies to frequencies further down in the distribution than Pareto's α , it is still a measure that applies more to the upper part of the distribution than to the entire range. As an incidental part of the project, I have also incorporated Mary Jean Bowman's suggestion of fitting Gini's exponential function to the lower part of the distribution, which she called the "reversal of the Gini curve," as a corresponding measure of inequality emphasizing the lower quarter or half of the distribution [1, p. 85]. This function was fitted to observations not exceeding 27½ percent, and 45 percent, respectively, of recipients cumulated from the bottom. The values for the exponential coefficients, together with the cumulated frequencies to which each was fitted, are shown in the first eight columns of the table. It should be noted that, for the upper part of the distribution, higher values for B_2 correspond with greater inequality; while for the lower part of the distribution, higher values for B_2 imply less inequality.

Finally, from the composite function repre-

senting the Lorenz curve for each distribution, it was an easy matter to compute the Gini concentration ratio—the ratio of the area lying between the line of perfect equality and the Lorenz curve, to the area lying under the line of perfect equality—from the integral of the function. The values are shown in the next to the last column of Table 6. Since they have been computed from smooth curves rather than linear segments, they exceed somewhat values computed by others. For example, for CPS families and unrelated individuals combined, they average about .7 percentage points greater for the CPS than those computed by Paul Schultz [13, p. 79] and about 1.5 percentage points greater than those I computed several years ago by a method outlined by Morgan [2, p. xiii].

Empirical estimates of size distributions used for this study were drawn from the old OBE series [11], the CPS [5], the BLS's Survey of Consumer Expenditure [7], and IRS's *Statistics of Income* [9]; excluded were the Survey of Consumer Finances and the 1950 and 1960 Census, as well as distributions for single or two contiguous

TABLE 4

SHARE OF INCOME FOR EACH 5 PERCENT OF RECIPIENTS, FROM
LOWEST TO HIGHEST, 1, 2, . . . , 20, AND TOP 1 PERCENT

	1ST	2ND	3RD	4TH	5TH	6TH	7TH	8TH	9TH	10TH	11TH	12TH	13TH	14TH	15TH	16TH	17TH	18TH	19TH	20TH	TOP 1
CPS: FAMILIES + UNRELATED INDIVIDUALS: MONEY INCOME BEFORE TAX																					
1944 CPS	0.84	1.04	1.45	1.86	2.27	2.68	3.10	3.53	3.96	4.41	4.86	5.33	5.82	6.32	6.91	7.62	8.63	10.46	18.90	6.80	
1945 CPS	0.84	1.22	1.70	2.15	2.51	2.97	3.35	3.72	4.10	4.48	4.87	5.28	5.72	6.19	6.86	7.60	8.54	10.21	17.62	6.09	
1947 CPS	0.72	1.11	1.57	2.01	2.43	2.83	3.22	3.61	4.00	4.39	4.80	5.22	5.66	6.14	6.81	7.58	8.59	10.42	18.91	6.83	
1948 CPS	0.71	1.10	1.58	2.03	2.46	2.88	3.29	3.69	4.09	4.48	4.88	5.28	5.69	6.12	6.82	7.48	8.48	10.28	18.65	6.73	
1949 CPS	0.64	1.05	1.52	1.97	2.41	2.84	3.25	3.67	4.08	4.49	4.91	5.32	5.75	6.19	6.94	7.65	8.63	10.39	18.32	6.45	
1950 CPS	0.58	1.00	1.47	1.93	2.39	2.84	3.28	3.71	4.13	4.54	4.94	5.33	5.71	6.08	6.90	7.65	8.63	10.41	18.49	6.55	
1951 CPS	0.61	1.09	1.60	2.08	2.54	2.98	3.40	3.82	4.22	4.61	4.99	5.37	5.74	6.23	6.87	7.50	8.45	10.15	17.77	6.22	
1952 CPS	0.63	1.13	1.63	2.10	2.53	2.95	3.35	3.74	4.14	4.53	4.94	5.37	5.82	6.31	6.86	7.54	8.48	10.18	17.76	6.20	
1954 CPS	0.54	0.97	1.44	1.90	2.35	2.80	3.24	3.68	4.12	4.56	5.00	5.45	5.90	6.36	6.94	7.79	8.73	10.42	17.81	6.11	
1955 CPS	0.62	1.03	1.51	1.97	2.43	2.87	3.31	3.75	4.18	4.60	5.03	5.46	5.89	6.33	7.03	7.64	8.61	10.26	17.43	5.95	
1956 CPS	0.63	1.07	1.56	2.03	2.48	2.92	3.35	3.78	4.20	4.62	5.05	5.48	5.92	6.37	6.99	7.65	8.56	10.18	17.18	5.83	
1957 CPS	0.68	0.94	1.39	1.85	2.33	2.82	3.32	3.81	4.30	4.77	5.22	5.65	6.05	6.42	7.08	7.73	8.62	10.18	16.82	5.60	
1958 CPS	0.66	1.06	1.52	1.99	2.44	2.88	3.33	3.76	4.20	4.63	5.06	5.49	5.93	6.36	7.03	7.84	8.72	10.28	16.83	5.57	
1959 CPS	0.65	1.04	1.50	1.95	2.39	2.83	3.27	3.71	4.14	4.58	5.03	5.48	5.94	6.41	7.00	7.72	8.64	10.29	17.43	5.93	
1960 CPS	0.64	1.04	1.50	1.95	2.40	2.84	3.27	3.70	4.13	4.57	5.00	5.44	5.89	6.35	7.04	7.67	8.61	10.29	17.66	6.08	
1961 CPS	0.67	1.02	1.45	1.89	2.32	2.75	3.18	3.61	4.04	4.48	4.92	5.37	5.83	6.30	7.00	7.76	8.73	10.47	18.18	6.33	
1962 CPS	0.71	1.08	1.53	1.97	2.40	2.82	3.24	3.66	4.08	4.52	4.97	5.45	5.95	6.48	7.05	7.81	8.73	10.34	17.21	5.77	
1963 CPS	0.77	1.06	1.51	1.95	2.39	2.83	3.27	3.70	4.14	4.58	5.03	5.49	5.95	6.42	7.14	7.76	8.66	10.26	17.09	5.73	
1964 CPS	0.77	1.08	1.52	1.95	2.38	2.82	3.25	3.68	4.12	4.56	5.01	5.46	5.93	6.40	7.12	7.81	8.71	10.31	17.13	5.74	
1965 CPS	0.85	1.10	1.54	1.97	2.40	2.83	3.26	3.68	4.12	4.55	5.00	5.45	5.92	6.40	7.07	7.74	8.65	10.26	17.19	5.80	
1966 CPS	0.94	1.12	1.56	1.99	2.42	2.85	3.28	3.71	4.14	4.58	5.01	5.45	5.90	6.35	7.09	7.71	8.61	10.21	17.06	5.76	
1967 CPS	0.94	1.14	1.58	2.00	2.43	2.84	3.26	3.67	4.09	4.52	4.96	5.41	5.88	6.37	7.00	7.64	8.56	10.22	17.50	6.01	
1968 CPS	1.03	1.20	1.64	2.07	2.49	2.91	3.33	3.74	4.16	4.58	5.00	5.44	5.89	6.49	6.94	7.54	8.44	10.06	17.05	5.81	
OBE: FAMILIES + UNATTACHED INDIVIDUALS: FAMILY PERSONAL INCOME																					
1947 OBE	1.51	1.50	1.92	2.30	2.64	2.95	3.25	3.54	3.84	4.14	4.47	4.82	5.22	5.66	6.22	6.93	8.05	10.11	20.94	8.39	
1950 OBE	1.40	1.43	1.84	2.22	2.57	2.90	3.21	3.52	3.83	4.15	4.49	4.87	5.27	5.71	6.12	6.85	7.97	10.09	21.55	8.83	
1955 OBE	1.43	1.45	1.94	2.33	2.69	3.03	3.34	3.65	3.95	4.27	4.60	4.96	5.35	5.79	6.28	6.86	7.90	9.89	20.24	8.03	
1960 OBE	1.28	1.40	1.84	2.23	2.59	2.93	3.26	3.57	3.90	4.24	4.60	4.99	5.43	5.92	6.46	7.17	8.28	10.21	19.69	7.46	
1961 OBE	1.34	1.40	1.82	2.21	2.56	2.90	3.22	3.55	3.88	4.22	4.58	4.98	5.41	5.90	6.44	7.20	8.33	10.27	19.80	7.50	
BLS: SPENDING UNITS: MONEY INCOME AFTER TAX																					
1950 BLS	2.14	1.98	2.47	2.88	3.25	3.56	3.85	4.11	4.36	4.61	4.89	5.19	5.52	5.92	6.35	6.82	7.63	9.09	15.40	5.25	
1960 BLS	2.10	1.84	2.22	2.69	3.06	3.40	3.73	4.04	4.35	4.66	4.99	5.35	5.74	6.18	6.66	7.03	7.79	9.17	14.92	4.90	
IRS: ALL TAX RETURNS (ALL) 6 JOINT RETURNS (JNT): ADJUSTED GROSS INCOME																					
1948 ALL	1.27	1.36	1.78	2.18	2.54	2.89	3.22	3.54	3.86	4.18	4.52	4.87	5.25	5.65	6.09	6.45	7.54	9.77	23.06	10.15	
1950 ALL	1.23	1.40	1.84	2.24	2.60	2.94	3.25	3.56	3.86	4.18	4.52	4.89	5.30	5.75	6.00	6.63	7.77	9.94	22.11	9.34	
1955 ALL	1.02	1.18	1.59	1.99	2.38	2.76	3.14	3.51	3.90	4.29	4.70	5.13	5.58	6.03	6.43	7.15	8.25	10.28	20.68	8.10	
1960 ALL	0.88	1.04	1.44	1.84	2.25	2.66	3.07	3.49	3.91	4.34	4.78	5.23	5.74	6.19	6.70	7.41	8.48	10.45	20.11	7.61	
1964 ALL	0.66	0.90	1.29	1.68	2.09	2.50	2.92	3.36	3.80	4.27	4.74	5.24	5.85	6.26	6.79	7.52	8.63	10.68	20.85	7.98	
1966 ALL	0.69	0.83	1.19	1.57	1.98	2.41	2.86	3.33	3.81	4.31	4.83	5.35	5.89	6.44	6.97	7.69	8.77	10.76	20.32	7.58	
1967 ALL	0.76	0.83	1.17	1.55	1.95	2.38	2.82	3.29	3.77	4.27	4.77	5.29	5.81	6.33	6.90	7.63	8.75	10.60	20.92	7.95	
1955 JNT	1.97	1.79	2.23	2.62	2.96	3.27	3.56	3.83	4.09	4.36	4.64	4.95	5.29	5.65	5.89	6.56	7.58	9.48	15.30	7.63	
1960 JNT	1.72	1.70	2.15	2.55	2.91	3.25	3.55	3.85	4.14	4.43	4.74	5.06	5.41	5.80	6.15	6.69	7.68	9.51	18.61	7.13	
1964 JNT	1.84	1.74	2.19	2.59	2.95	3.28	3.58	3.87	4.15	4.44	4.73	5.05	5.40	5.79	6.16	6.73	7.70	9.50	18.25	6.93	
1966 JNT	2.15	1.87	2.29	2.66	2.98	3.27	3.53	3.79	4.05	4.32	4.62	4.95	5.32	5.76	6.26	6.84	7.49	9.07	18.79	7.53	
1967 JNT	1.86	1.77	2.23	2.62	2.97	3.28	3.57	3.84	4.10	4.38	4.67	4.99	5.35	5.76	6.23	6.77	7.39	9.05	19.18	7.82	

years—the 1962 Survey of Financial Characteristics of Consumers and the 1965 and 1966 Surveys of Economic Opportunity. The OBE series was run separately for families⁶ and for consumer units for selected years: 1946, 1947, 1950, 1955, 1960, and 1961. Virtually all available years of the CPS were used, for families,⁶ and the sum of families and unrelated individuals. The BLS for 1950 is available only for urban consumer units; the 1960–61 data used here are therefore restricted to the urban sample. The IRS data on individual returns are for selected years only: 1955, 1960, 1964, 1966, and 1967 for joint returns; 1948 and 1950 in addition to these years for all returns (joint, separate, head of household, surviving spouse, and single returns).

It may, perhaps, be unnecessary to warn the reader that size distributions from different sources are simply not comparable with each

other. To determine temporal changes in inequality, a given source at one point in time can be compared only with that same—not some other—source at a different point in time. Size distributions differ in at least four important ways: the concept of recipient units, adjustment for part-time recipient units, definition and coverage of income.⁷

The tax return is obviously not a particularly meaningful recipient unit concept. It refers to earners or individual income recipients, not to families, and not all who receive income are required to file returns. An analogy does exist, to be sure, between joint returns and married couples, and between other heads, other relatives, and unrelated individuals on the one hand, and the remaining returns on the other. But it is a difficult job, and a job not done in *Statistics of Income*, to combine these returns into appropriate family units.

⁶ Tables for the family distributions have not been included in this paper for lack of space; they may be obtained by writing the author.

⁷ For more detailed discussions of these differences, see [4] and [12, Chap. 3].

TABLE 5

RELATIVE MEAN INCOME FOR EACH 5 PERCENT OF RECIPIENTS, FROM LOWEST TO HIGHEST

	1ST	2ND	3RD	4TH	5TH	6TH	7TH	8TH	9TH	10TH	11TH	12TH	13TH	14TH	15TH	16TH	17TH	18TH	19TH	20TH	TOP 1%
CPS: FAMILIES + UNRELATED INDIVIDUALS: MONEY INCOME BEFORE TAX																					
1944 CPS	0.08	0.21	0.29	0.37	0.45	0.54	0.62	0.71	0.79	0.88	0.97	1.07	1.16	1.26	1.38	1.52	1.73	2.09	3.78	6.80	
1945 CPS	0.08	0.24	0.34	0.43	0.51	0.59	0.67	0.74	0.82	0.90	0.97	1.06	1.14	1.24	1.37	1.52	1.71	2.04	3.52	6.09	
1947 CPS	0.07	0.22	0.31	0.40	0.49	0.57	0.64	0.72	0.80	0.88	0.96	1.04	1.13	1.23	1.36	1.52	1.72	2.08	3.78	6.83	
1948 CPS	0.07	0.22	0.32	0.41	0.49	0.58	0.66	0.74	0.82	0.90	0.98	1.06	1.14	1.22	1.36	1.50	1.70	2.06	3.73	6.73	
1949 CPS	0.06	0.21	0.30	0.39	0.48	0.57	0.65	0.73	0.82	0.90	0.98	1.06	1.15	1.24	1.39	1.53	1.73	2.08	3.66	6.45	
1950 CPS	0.06	0.20	0.29	0.39	0.48	0.57	0.66	0.74	0.83	0.91	0.99	1.07	1.14	1.22	1.38	1.53	1.73	2.08	3.70	6.55	
1951 CPS	0.06	0.22	0.32	0.42	0.51	0.60	0.68	0.76	0.84	0.92	1.00	1.07	1.15	1.25	1.37	1.50	1.69	2.03	3.55	6.22	
1952 CPS	0.06	0.23	0.33	0.42	0.51	0.59	0.67	0.75	0.83	0.91	0.99	1.07	1.17	1.26	1.37	1.51	1.70	2.04	3.55	6.20	
1954 CPS	0.05	0.19	0.29	0.38	0.47	0.56	0.65	0.74	0.82	0.91	1.00	1.09	1.18	1.27	1.39	1.56	1.75	2.08	3.56	6.11	
1955 CPS	0.06	0.21	0.30	0.39	0.49	0.57	0.66	0.75	0.84	0.92	1.01	1.09	1.18	1.27	1.41	1.54	1.72	2.05	3.49	5.95	
1956 CPS	0.06	0.21	0.31	0.41	0.50	0.58	0.67	0.76	0.84	0.92	1.01	1.10	1.18	1.27	1.40	1.53	1.71	2.04	3.44	5.83	
1957 CPS	0.07	0.19	0.28	0.37	0.47	0.56	0.66	0.76	0.86	0.95	1.04	1.13	1.21	1.28	1.42	1.55	1.72	2.04	3.36	5.60	
1958 CPS	0.07	0.21	0.30	0.40	0.49	0.58	0.67	0.76	0.84	0.93	1.01	1.10	1.19	1.27	1.41	1.57	1.74	2.06	3.49	5.93	
1959 CPS	0.06	0.21	0.30	0.39	0.48	0.57	0.65	0.74	0.82	0.92	1.01	1.10	1.19	1.28	1.40	1.54	1.73	2.06	3.43	5.57	
1960 CPS	0.06	0.21	0.30	0.39	0.48	0.57	0.65	0.74	0.83	0.91	1.00	1.09	1.18	1.27	1.41	1.53	1.72	2.06	3.53	6.08	
1961 CPS	0.07	0.20	0.29	0.38	0.46	0.55	0.64	0.72	0.81	0.90	0.98	1.07	1.17	1.26	1.40	1.55	1.75	2.09	3.64	6.33	
1962 CPS	0.07	0.22	0.31	0.39	0.48	0.56	0.65	0.73	0.82	0.90	0.99	1.09	1.19	1.30	1.41	1.56	1.75	2.07	3.44	5.77	
1963 CPS	0.08	0.21	0.30	0.39	0.48	0.57	0.65	0.74	0.83	0.92	1.01	1.10	1.19	1.28	1.43	1.55	1.73	2.05	3.42	5.73	
1964 CPS	0.08	0.22	0.30	0.39	0.48	0.56	0.65	0.74	0.82	0.91	1.00	1.09	1.19	1.28	1.42	1.56	1.74	2.06	3.43	5.74	
1965 CPS	0.09	0.22	0.31	0.39	0.48	0.57	0.65	0.74	0.82	0.91	1.00	1.09	1.18	1.28	1.41	1.55	1.73	2.05	3.44	5.80	
1966 CPS	0.09	0.22	0.31	0.40	0.48	0.57	0.66	0.74	0.83	0.92	1.00	1.09	1.18	1.27	1.42	1.54	1.72	2.04	3.42	5.76	
1967 CPS	0.09	0.23	0.32	0.40	0.49	0.57	0.65	0.73	0.82	0.90	0.99	1.08	1.18	1.27	1.40	1.53	1.71	2.04	3.50	6.01	
1968 CPS	0.10	0.24	0.33	0.41	0.50	0.58	0.67	0.75	0.83	0.92	1.00	1.09	1.18	1.30	1.39	1.51	1.69	2.01	3.41	5.81	
OBE: FAMILIES + UNATTACHED INDIVIDUALS: FAMILY PERSONAL INCOME																					
1947 OBE	0.15	0.30	0.38	0.46	0.53	0.59	0.65	0.71	0.77	0.83	0.89	0.96	1.04	1.13	1.24	1.39	1.61	2.02	4.19	8.39	
1950 OBE	0.14	0.29	0.37	0.44	0.51	0.58	0.64	0.70	0.77	0.83	0.90	0.97	1.05	1.14	1.22	1.37	1.59	2.02	4.31	8.83	
1955 OBE	0.14	0.30	0.39	0.47	0.54	0.61	0.67	0.73	0.79	0.85	0.92	0.99	1.07	1.16	1.26	1.37	1.58	1.98	4.05	8.03	
1960 OBE	0.13	0.28	0.37	0.45	0.52	0.59	0.65	0.71	0.78	0.85	0.92	1.00	1.09	1.18	1.29	1.43	1.66	2.04	3.94	7.46	
1961 OBE	0.13	0.28	0.36	0.44	0.51	0.58	0.64	0.71	0.78	0.84	0.92	1.00	1.08	1.18	1.29	1.44	1.67	2.05	3.96	7.50	
BLS: SPENDING UNITS: MONEY INCOME AFTER TAX																					
1950 BLS	0.21	0.40	0.49	0.58	0.65	0.71	0.77	0.82	0.87	0.92	0.96	1.04	1.10	1.18	1.27	1.36	1.53	1.82	3.08	5.25	
1960 BLS	0.21	0.37	0.46	0.54	0.61	0.68	0.75	0.81	0.87	0.93	1.00	1.07	1.15	1.24	1.33	1.41	1.56	1.83	2.98	4.90	
IRS: ALL TAX RETURNS (ALL) & JOINT RETURNS (JNT): ADJUSTED GROSS INCOME																					
1948 ALL	0.13	0.27	0.36	0.44	0.51	0.58	0.64	0.71	0.77	0.84	0.90	0.97	1.05	1.13	1.22	1.29	1.51	1.95	4.61	10.15	
1950 ALL	0.12	0.28	0.37	0.45	0.52	0.59	0.65	0.71	0.77	0.84	0.90	0.98	1.06	1.15	1.20	1.33	1.55	1.99	4.42	9.34	
1955 ALL	0.10	0.24	0.32	0.40	0.48	0.55	0.63	0.70	0.78	0.86	0.94	1.03	1.12	1.21	1.29	1.43	1.65	2.06	4.14	8.10	
1960 ALL	0.09	0.21	0.29	0.37	0.45	0.53	0.61	0.70	0.78	0.87	0.96	1.05	1.15	1.24	1.34	1.48	1.70	2.09	4.02	7.61	
1964 ALL	0.07	0.18	0.26	0.34	0.42	0.50	0.58	0.67	0.76	0.85	0.95	1.05	1.17	1.25	1.36	1.50	1.73	2.14	4.17	7.98	
1966 ALL	0.07	0.17	0.24	0.31	0.40	0.48	0.57	0.67	0.76	0.86	0.97	1.07	1.18	1.29	1.39	1.54	1.75	2.15	4.07	7.58	
1967 ALL	0.08	0.17	0.23	0.31	0.39	0.48	0.56	0.66	0.75	0.85	0.95	1.06	1.16	1.27	1.38	1.53	1.75	2.16	4.18	7.95	
1955 JNT	0.20	0.36	0.45	0.52	0.59	0.65	0.71	0.77	0.82	0.87	0.93	0.99	1.06	1.13	1.18	1.31	1.52	1.90	3.86	7.63	
1960 JNT	0.18	0.34	0.43	0.51	0.58	0.65	0.71	0.77	0.83	0.89	0.95	1.01	1.08	1.16	1.23	1.34	1.54	1.90	3.72	7.13	
1964 JNT	0.18	0.35	0.44	0.52	0.59	0.66	0.72	0.77	0.83	0.89	0.95	1.01	1.08	1.16	1.23	1.35	1.54	1.90	3.66	6.93	
1966 JNT	0.22	0.37	0.46	0.53	0.60	0.67	0.73	0.79	0.81	0.86	0.92	0.99	1.06	1.15	1.25	1.37	1.50	1.81	3.76	7.53	
1967 JNT	0.19	0.35	0.45	0.52	0.59	0.66	0.71	0.77	0.82	0.88	0.93	1.00	1.07	1.15	1.25	1.35	1.48	1.81	3.84	7.82	

OBE and the CPS are both based on essentially the same recipient unit concept: families and unrelated individuals, while the BLS (and the Survey of Consumer Finances through 1962 as well) employs the spending unit. Differences on this score, however, are relatively minor compared with the adjustment, or absence thereof, for part-year income recipient units, particularly in comparisons of the CPS and the BLS. The CPS is essentially a distribution of previous year's income reported by recipient units defined as of the date of the survey (March or April of the following year), even though the unit may have had a different family status in the previous year, or if a head, might not have been head of a family for all of that year. BLS, on the other hand, includes only full-year unrelated individuals, and other relatives only for the period they are living in the family. The major difference is in the number of unrelated individuals with low incomes included in the CPS; as a result, the CPS distributions for consumer units show considerably more inequality, particularly at the bottom of the distribution. It would require a separate paper to examine these problems more thoroughly; my own particular view is that, while the CPS concepts are well suited to demographic and

population work, and the BLS, to consumption expenditure estimates, neither is necessarily appropriate for income size work. We need data on income size which are not derived simply as a by-product of some other purpose, whether it be for tax collection, demography, or consumer expenditure.

Virtually all of the tabulations of tax returns published in *Statistics of Income* are based on adjusted gross income (AGI)—one of the least satisfactory income concepts one can think of for size distributions. The omission of state and local bond interest and most transfer income and the inclusion of capital and other types of realized gains and losses are among the most serious, although not the only, shortcomings. CPS and BLS, on the other hand, both use a similar concept of money income, although the BLS distributions used here are based on income after the deduction of personal taxes, whereas CPS incomes are before taxes. The major differences between money income as used in these field surveys and OBE's family personal income are the latter's exclusion of employee and self-employed social security contributions, and private pensions and transfers, and the inclusion of various items of imputed income.

TABLE 6
GINI CONCENTRATION, INTERSECTION DATA, AND CURVE FIT PARAMETERS

	1	2	3	4	5	6	7	8	9	10	11
CPS: FAMILIES + UNRELATED INDIVIDUALS: MONEY INCOME BEFORE TAX											
1944 CPS	24.7	1.574	34.2	1.587	34.4	0.502	45.0	0.502	0.436	77.1	-0.04
1945 CPS	26.3	1.516	35.6	1.549	28.5	0.491	40.6	0.500	0.411	77.0	-0.07
1947 CPS	26.2	1.579	33.5	1.606	35.8	0.481	35.8	0.481	0.430	76.9	-0.07
1948 CPS	28.7	1.578	36.6	1.602	33.3	0.474	43.5	0.482	0.424	74.8	-0.06
1949 CPS	27.3	1.543	35.0	1.570	35.0	0.470	45.2	0.476	0.428	76.0	-0.08
1950 CPS	24.2	1.552	30.9	1.571	32.8	0.458	41.8	0.462	0.431	75.9	-0.11
1951 CPS	29.6	1.534	37.6	1.556	28.5	0.448	44.1	0.465	0.416	73.1	-0.07
1952 CPS	27.3	1.530	33.6	1.540	34.7	0.461	42.0	0.469	0.416	78.8	-0.03
1954 CPS	21.2	1.505	31.4	1.527	28.1	0.444	40.5	0.454	0.429	78.4	-0.09
1955 CPS	24.0	1.498	34.8	1.520	31.9	0.452	44.8	0.464	0.420	75.6	-0.07
1956 CPS	27.8	1.490	39.7	1.528	29.3	0.470	40.6	0.473	0.415	76.2	-0.06
1957 CPS	29.3	1.464	41.9	1.509	29.0	0.513	45.2	0.474	0.418	75.9	-0.08
1958 CPS	22.2	1.454	31.3	1.481	28.5	0.472	44.5	0.477	0.416	77.5	-0.10
1959 CPS	26.0	1.494	35.4	1.522	32.0	0.472	42.1	0.476	0.422	77.5	-0.06
1960 CPS	28.4	1.510	37.8	1.541	30.6	0.470	45.6	0.476	0.423	74.9	-0.06
1961 CPS	23.0	1.524	30.7	1.544	30.3	0.476	44.9	0.481	0.432	76.8	-0.08
1962 CPS	25.2	1.473	32.7	1.500	28.8	0.486	42.7	0.489	0.421	80.5	-0.05
1963 CPS	28.0	1.474	35.9	1.499	27.6	0.485	44.9	0.488	0.418	75.3	-0.07
1964 CPS	23.6	1.471	30.0	1.484	31.1	0.488	43.8	0.491	0.419	76.3	-0.08
1965 CPS	26.3	1.480	33.2	1.501	29.1	0.498	41.0	0.499	0.417	76.4	-0.07
1966 CPS	24.6	1.479	30.5	1.490	30.7	0.520	38.0	0.514	0.413	75.1	-0.07
1967 CPS	28.3	1.506	34.2	1.521	28.8	0.524	43.1	0.514	0.416	75.5	-0.05
1968 CPS	33.1	1.495	39.4	1.515	32.1	0.529	38.9	0.525	0.406	69.7	-0.06
OBE: FAMILIES + UNATTACHED INDIVIDUALS: FAMILY PERSONAL INCOME											
1947 OBE	24.1	1.759	36.9	1.769	43.8	0.587	43.8	0.587	0.415	75.5	-0.03
1950 OBE	28.2	1.804	42.6	1.798	39.8	0.575	39.8	0.575	0.423	74.6	0.01
1955 OBE	19.4	1.742	32.7	1.723	41.2	0.574	41.2	0.574	0.407	81.8	-0.01
1960 OBE	16.8	1.659	31.3	1.671	32.9	0.557	43.9	0.561	0.415	82.6	-0.05
1961 OBE	17.6	1.659	32.3	1.678	32.3	0.568	43.1	0.568	0.417	82.0	-0.06
BLS: SPENDING UNITS: MONEY INCOME AFTER TAX											
1950 BLS	21.9	1.494	38.8	1.493	37.3	0.620	37.3	0.620	0.330	79.2	0.00
1960 BLS	41.3	1.446	54.4	1.485	32.4	0.630	45.6	0.622	0.338	80.4	0.02
IRS: ALL TAX RETURNS (ALL) & JOINT RETURNS (JNT): ADJUSTED GROSS INCOME											
1948 ALL	21.2	1.962	39.4	1.883	36.5	0.558	36.5	0.558	0.431	81.1	0.04
1950 ALL	26.3	1.867	45.0	1.828	33.2	0.546	33.2	0.546	0.425	76.3	0.06
1955 ALL	29.6	1.717	43.4	1.718	41.6	0.531	41.6	0.531	0.435	74.1	0.01
1960 ALL	23.0	1.656	31.8	1.652	35.0	0.517	35.0	0.517	0.443	66.8	-0.04
1964 ALL	25.0	1.676	32.4	1.672	32.0	0.486	41.4	0.483	0.461	65.0	-0.05
1966 ALL	24.8	1.631	31.2	1.631	30.7	0.495	39.0	0.489	0.464	70.7	-0.03
1967 ALL	27.3	1.665	33.5	1.669	29.2	0.516	44.8	0.492	0.468	74.5	-0.02
1955 JNT	29.4	1.735	43.4	1.706	31.3	0.618	39.2	0.619	0.373	74.5	0.03
1960 JNT	25.1	1.679	34.4	1.654	28.6	0.603	40.7	0.603	0.375	77.6	0.02
1964 JNT	23.1	1.658	30.3	1.645	29.9	0.605	39.5	0.606	0.370	77.5	0.01
1966 JNT	9.8	1.760	31.4	1.675	31.2	0.632	40.3	0.629	0.366	88.8	0.01
1967 JNT	11.7	1.793	35.3	1.681	28.8	0.605	37.4	0.607	0.372	88.2	-0.00

Key to column headings for Table 6: Coefficients (exponent) for Gini curve (upper part of distribution):

1. Percentile (<30%) to which Gini curve was fitted.
2. Coefficient for fitted curve.
3. Percentile to which Gini curve was fitted when one additional observation was included.
4. Corresponding coefficient for fitted curve.

Coefficients (exponent) for reversal of Gini curve (lower part of distribution):

5. Percentile (<27.5%) to which Gini curve was fitted. If >27.5%, not enough observations <27.5%, to fit curve.
6. Corresponding coefficient of fitted curve.
7. Percentile (<46%) to which curve was fitted.
8. Corresponding coefficient.
9. Gini concentration ratio.
10. Intersection point of polynomial function and exponential (Gini) function.
11. Relative mean income difference for percentile within which polynomial and exponential intersect.

While the inclusion of imputed income obviously serves to raise average income, its effect on inequality is far from clear. Low-income farm residents may gain a little from the imputation for food and fuel consumed on farms. Judging from the distribution of asset ownership by age and income size revealed by the 1962 Survey of Financial Characteristics of Consumers, which we are using at OBE in allocating certain types of imputed income by size, the lower-income aged may gain a little relatively from imputed rent, and the upper brackets from imputed interest and possibly imputed rent. If anything, there may be a small relative redistribution away from the middle-income groups towards the rich and the poor, implying an intersection of the Lorenz curves for money income and for money and imputed income combined.

Differences in the extent of income coverage and underreporting again may well be more important in their effects on inequality shown by the various data sources than are differences in income concepts. Only if undercoverage were the same at all income levels and for all income types could we ignore its effect on relative distribution. Only the OBE series accounts fully for all of the income that it is defined to cover. The CPS comes close to being a distribution of earnings plus social security payments: it accounts for only about a third to two-fifths of the relevant control totals for property income and not much more than half of other transfers excluding social security. IRS, on the other hand, does better on interest and dividends, especially the latter, but is particularly deficient in its coverage of rent and farm income. The IRS coverage, especially for rent and self-employment income, can be improved by corrections based on audit studies, such as the set of correction ratios for each income type that we have worked out at OBE, but such corrected data are unfortunately not available in *Statistics of Income* tabulations or to private research workers. It should also be noted that the shapes of the distributions of individual income types in the CPS and the IRS are quite different, quite apart from their failure to come up to the relevant control totals. This is particularly evident for self-employment income, which shows considerably more dispersion, and more negative incomes, in the tax return data than in the CPS.

A brief review of some of the results Daniel Radner and I presented at last year's meetings [3], but cast into the framework and the computer programs used for this paper, may throw some light on the effect of underreporting on relative

distribution. In that paper, we blew up each income type for each record, separately for families and unrelated individuals, by the ratio of the OBE control total for that type to the total amount actually reported in the CPS, and carried through the same procedure after substituting the tax return amount of earnings and property income for the CPS amount for those records to which we had assigned a tax return in our statistical matching procedure. I have combined the separate distributions given in that paper—for families plus subfamilies and for unrelated individuals—into distributions for consumer units. Table 7 summarizes the relative distributions for the same quantiles as used in Table 1. Note that the results of adjusting each distribution to the relevant control total separately for each income type is to raise the income shares of the lower part and the top tail for both the CPS and the tax return distributions. The CPS set of Lorenz curves intersect at the 65th, the tax return ones, at the 75th, percentile. Since the relative mean incomes of the CPS before and after inflation are equal at the 39th and 94th percentiles, and for the tax return distribution at the 42nd and the 96th percentiles, we can infer that the shares of the lowest 39 and 42 percent, and the top 6 and 4 percent, respectively, were raised, and the share of the group lying between those limits was reduced, by our inflation procedure.

Emphasizing the noncomparability of different size distributions may be belaboring the obvious. It should be recognized, however, that size distributions drawn from the same source are themselves not necessarily comparable over time. Temporal differences in coverage of income or of recipients, possible changes in concepts, and changes in procedure, e.g., enumerator quality, editing, processing, programming—all take their toll. *Statistics of Income* must contend with changes in tax laws, which affect the definitions of income types or the returns that must be filed. The extension of the split income privilege to all married taxpayers in 1948 certainly affects the comparability of tax return distributions for 1948 and subsequent years with those for earlier years; years prior to 1948 for IRS have therefore been excluded from this study. There has also been a gradual improvement in compliance. One of the most striking improvements occurred when reporting by financial and other institutions to IRS of interest and dividends paid to individuals was first instituted. In the two years between 1961 and 1963 the percent of returns reporting interest rose from 16 to 33, and dividends, from 11 to 15.

In addition to sampling error, the CPS is sub-

TABLE 7

DISTRIBUTION OF MONEY INCOME IN 1964 FOR CONSUMER UNITS,* BASED ON THE
CPS AND ON TAX RETURN DATA, BEFORE AND AFTER ADJUSTMENT TO
OBE MONEY INCOME CONTROL TOTAL

A. Percentile (cumulated)								
	20	40	60	80	95	99	100	Gini Concen- tration Ratio
<i>Current Population Survey</i>								
Before adjustment.....	3.2	13.5	30.8	55.8	82.8	94.3	100	.423
After adjustment.....	3.5	14.2	30.9	55.2	81.6	93.5	100	.424
<i>Adjusted by tax returns</i>								
Before adjustment.....	3.0	13.2	30.1	54.6	81.1	93.2	100	.437
After adjustment.....	3.6	14.1	30.8	54.2	80.1	92.4	100	.431

B. Income share of quantile							
	Quintiles		Middle	Fourth	Next 15%	Top 5%	Top 1%
	Lowest	Second					
<i>Current Population Survey</i>							
Before adjustment.....	3.2	10.3	17.3	25.0	27.0	11.5	5.7
After adjustment.....	3.5	10.7	16.7	24.3	26.4	11.9	6.5
<i>Adjusted by tax returns</i>							
Before adjustment.....	3.0	10.2	16.9	24.5	26.5	12.1	6.8
After adjustment.....	3.6	10.5	16.7	23.4	25.9	12.3	7.6

C. Relative mean income of quantile							
<i>Current Population Survey</i>							
Before adjustment.....	.16	.515	.865	1.25	1.93	2.30	5.70
After adjustment.....	.175	.535	.835	1.215	1.78	2.38	6.50
<i>Adjusted by tax returns</i>							
Before adjustment.....	.15	.51	.845	1.225	1.79	2.42	6.80
After adjustment.....	.18	.525	.835	1.17	1.78	2.46	7.60

* Subfamilies counted separately.

SOURCE: Office of Business Economics. For details, see [3].

ject to a whole host of variations in such things as interviewing, editing, and processing, that it is hazardous to draw any firm conclusions about year-to-year changes in CPS distributions. This is particularly true of the unrelated individual distribution, which is quite sensitive to the method used to convert weights for individual persons in the CPS to weights for consumer units. The comparability of secular changes may also be affected by improvements in the survey and its processing, such as the introduction by Census for the 1961 income year of an income assignment method to nonrespondents to the income questions, and the adding of one more "other income" question in

1966.⁸ The proportion of nonrespondents has risen somewhat over time; allocation procedures show that nonrespondents have demographic and economic characteristics (e.g., a higher proportion of self-employed) that are associated with higher incomes. Further, the amount of income reported is a function of the number of questions asked on

⁸ For income year 1966, Census has also published distributions before and after the introduction of revised editing and nonrespondent income allocation procedures for the March, 1967, CPS. For this year, therefore, the difference in CPS size distribution resulting from the difference in methods can actually be computed [5, No. 59, pp. 16-19].

different income types. This is not meant to be a criticism of the changes Census has introduced; indeed, we may need more rather than fewer. Certainly we do not want to be stuck with obsolete methods just for the sake of preserving historical comparability of CPS size distributions.

The OBE series, insofar as it used these and other basic data sources and did not or could not correct for such temporal changes in comparability, would involve the same problem. But the estimating method itself, while undoubtedly giving the best estimate of size distribution in any one year, was perhaps biased toward stability in relative distribution over time. The same bench mark estimates of the relative distribution of income among unattached individuals and among farm operator families, and the same pattern of combining tax returns into nonfarm family units (by use of a 1950 Census-IRS matching study) were used from 1950 on. The relative distribution could change, therefore, only if either the tax return distribution changed or the weights of the three groups or their mean incomes changed relative to one another. Indeed, the obsolete nature of these bench marks finally led OBE to discontinue the series after publishing estimates for 1962. I should add, as a footnote, perhaps, that the different methodology employed in the new OBE series may make it difficult to draw inferences about changes in distribution employing both the new and the old series.

Keeping in mind these qualifications on the temporal comparisons of size distributions, it may be useful to summarize some of the results which show, for certain of the distributions in the postwar period, the intersection of Lorenz curves which the techniques outlined in this paper permit us to pinpoint. The BLS after tax money income distributions of spending units for 1950 and 1960 indicate, in a mild way, just this kind of pattern: the 1960 Lorenz curve intersects the 1950 one from below, at the 80th percentile. The relative mean incomes of the two distributions are equal at the 49th and 96th percentiles, giving us the precise points for determining the size groups which gained or lost. The relative mean income of the bottom 49 percent fell by 4 percent; that of the group comprising the 50th to the 95th percentiles rose by 3 percent; that of the top 4 percent fell by 4 percent. Smaller groups, of course, fared differently: the share of the group comprising the 16th to 20th percentiles, for example, fell by over 7 percent and that of the percentile corresponding to the intersection point rose by 5 percent.

While the technique developed in this paper does not really alter the conclusions drawn earlier

on changes shown by the OBE series for families and unattached individuals from the immediate postwar years to 1960-61, it does permit us to pinpoint the percentile at which the 1947 and 1961 Lorenz curves intersect—the 70th—and the quantile groups which gained or lost. To be exact, the relative mean of the bottom 42 percent fell from .42 to .40 or by 4 percent and that of the top 4 percent fell from 4.61 to 4.33 or by 6 percent. The mean of the middle 54 percent, on the other hand, rose from 1.19 to 1.22 or by 3 percent. Again, other more narrowly defined groups could have gained or lost more; the share of the top 1 percent, for example, fell by over 10 percent.

As noted earlier, the intersection phenomenon is quite evident in the distribution of all tax returns for 1948 and 1967. Similarly, the IRS distribution for 1960 cuts the 1948 one from below, with the intersection occurring at the 80th percentile. The relative mean income of the lowest 46 and the top 3 percent of tax return units fell by 11 and 17 percent, respectively, between these two years, while that of the middle 51 percent rose by 8 percent. The 1967 Lorenz curve, on the other hand, lies entirely below the 1960 one; the Gini concentration ratio increased by 2.6 percentage points in this seven-year period, compared with an increase of 1.1 percentage points between 1950 and 1960.

These differences can be detected but are less evident for the joint return data. Between 1955 and 1960, using the same method to locate the precise limits of the groups who gained or lost relative to the average, we find that the relative mean income of the first through 38th percentiles fell by 3 percent, that of the 39th through the 94th percentiles rose by 2 percent, and that of the top 6 percent fell by 3 percent. On the other hand, between 1960 and 1967, the share of the lower 78 percent of married couples filing joint returns was unchanged. The redistribution that did occur was within the top 22 percent. The relative mean income of the group encompassing the 78th to the 99th percentiles fell by 2 percent; that of the top 1 percent rose by 9 percent.

Because of some year-to-year fluctuations in income shares of quantile groups in the CPS, the existence of an intersection phenomena for earlier and later years in this source and the extent of gain or loss for different quantile groups if an intersection occurs is to some extent dependent on the years selected for comparison. It should also be noted that the CPS, in contrast to the tax return data, suggests an overall reduction in inequality from the late 1940's (1947-49) to the late 1960's (1965-68), without evidence of any

Lorenz curve intersection. The Gini concentration ratio, therefore, gives an unambiguous measure of the reduction; it fell from an average of 42.75 for the earlier years to an average of 41.31 for the later ones, a decline of 3.4 percent.

The intersection phenomenon in the CPS is evident, however, when we compare the late 1950's and early 1960's with the 1947-49 period. If, as in the previous comparisons the OBE, we take 1947 and 1961, we find that the curves intersect at the 70th percentile. The relative mean income of the bottom 43 percent fell from .372 to .358; that of the top 5 percent, from 3.78 to 3.64—both by 4 percent. The middle 52 percent experienced a rise in its mean from 1.25 to 1.28, or by 2 per-

cent. These changes confirm the pattern suggested by the other sources for these years.

Nothing has been said about the consistency of these findings with the movement of the Gini and the Gini reversal coefficients over the years compared. In all cases, the Gini coefficients, which, it will be recalled, are relevant to the upper part of the distribution, change in the right direction—they decline for all comparisons where the relative mean income of the top group also declines. The Gini reversal (relevant to the changes in inequality at the bottom) also shows the appropriate direction of change, with the exception of the BLS and the CPS comparisons. Because the bottom group encompasses such a large part of the

TABLE 8
RELATIVE MEAN INCOMES AND PERCENT CHANGE IN MEASURES OF INEQUALITY,
FOR SELECTED SERIES AND YEARS

	Upper Limit of Quan- tile	Size of Quan- tile	Relative Mean Income			Percent Change in	
			Initial Year	End Year	Percent Change	Gini Coeffi- cient*	Gini Con- centra- tion Ratio
<i>BLS: 1950 to 1960</i>							
Lower.....	49	49	.565	.565	-3.9	+ .3	+2.5
Middle.....	96	47	1.26	1.29	+2.7		
Upper.....	100	4	3.32	3.20	-3.6	-3.2	
<i>OBE: 1947 to 1961</i>							
Lower.....	42	42	.415	.401	-3.5	-3.2	+.4
Middle.....	96	54	1.19	1.22	+2.7		
Upper.....	100	4	4.61	4.33	-6.1	-5.7	
<i>CPS: 1947 to 1961</i>							
Lower.....	43	43	.372	.358	-3.8	0	+.4
Middle.....	95	52	1.25	1.28	+2.0		
Upper.....	100	5	3.78	3.64	-3.9	-3.5	
<i>IRS All: 1948 to 1960</i>							
Lower.....	46	46	.42	.38	-10.8	-7.4	+2.6
Middle.....	97	51	1.23	1.33	+8.1		
Upper.....	100	3	5.92	4.92	-16.9	-15.6	
<i>IRS Joint: 1955 to 1960</i>							
Lower.....	38	38	.46	.45	-2.8	-2.6	+.5
Middle.....	94	56	1.09	1.11	+2.0		
Upper.....	100	6	3.57	3.46	-3.2	-4.4	

* For the lower group, percent change in the exponent for reversal of Gini curve fitted to observations comprising bottom 46 percent of recipients or less; for the upper group, percent change in exponent of Gini curve fitted to observations comprising top 30 percent of recipients or less.

SOURCE: Derived from Tables 2 thru 6 or computer printout underlying them.

distribution (from 40 to 50 percent of the recipients in most cases) to which the reversal function was fit, it may well be that it is not as sensitive a measure of change at the bottom as is the relative mean income technique which has been emphasized in this paper.

The foregoing comparisons of intersection phenomena for the different series are summarized in Table 8. The first two columns give the percentile limits and sizes of the bottom, middle, and top groups; the next three, the size and percent change in relative mean incomes; the last three, the percent change in the three measures of inequality: the Gini coefficients for the top and bottom of the distribution and the Gini concentration ratio.

This paper has attempted several things. First, it has tried to provide a method for comparing different size distributions, or size distributions from the same source for different years, that emphasizes changes in distribution among component recipient groups (when classified by income size) rather than overall changes in inequality as represented, say, by the Gini concentration ratio. It does, of course, stay within the Lorenz curve framework, which ranks recipients in a given year by that year's income. To say that the relative mean income of the top 1 percent fell by x percent is obviously not to say that a specified group of recipients experienced that x percent fall, for the composition of the top 1 percent will not be the same from year to year. Second, it has applied the method to a number of different size distributions from different sources for selected war and postwar years and developed a means of interpolating the entire Lorenz curve for a given distribution from a knowledge of a limited number of points.

Finally, it has attempted to draw certain tentative conclusions about changes in the size distribution of income from the immediate postwar years to the 1960's from evidence drawn from a number of different distributions. This evidence points to a gain by the middle and upper part of the distribution, relative to the lower groups and the upper tail. The bottom group seems to encompass the first through the 40th or 50th percentiles; the middle group, from the 40th or 50th percentiles to the 94th or 96th percentile; the top group, the top 4 to 6 percent. The Lorenz curves for the early postwar years and the 1960's would appear to intersect between the 70th and the 80th percentiles. The changes in relative mean incomes are, of course, quite small, but they seem to be present in enough of the sources to suggest that these changes were characteristic of the actual

world rather than simply reflecting deficiencies in our data sources, although I would not want in any way to minimize those deficiencies.

REFERENCES

1. MARY JEAN BOWMAN, "A Graphical Analysis of Personal Income Distribution in the U.S." in *Readings in the Theory of Income Distribution* (Blakiston, 1951), pp. 72-99.
2. EDWARD C. BUDD, ed., *Inequality and Poverty* (W. W. Norton, 1967).
3. EDWARD C. BUDD AND DANIEL B. RADNER, "The OBE Size Distribution Series: Methods and Tentative Results for 1964," *A.E.R.*, May, 1969, pp. 435-49.
4. Bureau of the Budget, *Family Income Distribution Statistics Published by Federal Agencies*, Statistical Evaluation Report No. 5, Dec., 1964.
5. Bureau of the Census, *Current Population Reports*, Series P-60 (Consumer Income) Nos. 1-7, 9, 12, 15, 20, 24, 27, 33, 35, 37, 39, 41, 43, 47, 51, 53, 59.
6. Bureau of the Census, *Trends in the Income of Families and Persons in the U.S., 1947-1964*, Technical Paper No. 17 (U.S. Government Printing Office, 1967).
7. Bureau of Labor Statistics, *Study of Consumer Expenditure, Incomes, and Savings*, 1950 (Univ. of Pennsylvania Press, 1956), Vol. I, p. 2, and Survey of Consumer Expenditure, Reports 237-93, No. 1, May, 1966.
8. CORRADO GINI, "On the Measure of Concentration with Special Reference to Income and Wealth," abstracts of Papers Presented at the Cowles Commission Research Conference on Economics and Statistics (Colorado College Publication, 1936).
9. Internal Revenue Service, *Statistics of Income: Individual Income Tax Returns*, for 1948, 1950, 1955, 1960, 1964, 1966, and 1967 (U.S. Government Printing Office).
10. HERMAN P. MILLER, *Income Distribution in the U.S.*, 1960 Census Monograph (U.S. Government Printing Office, 1966).
11. Office of Business Economics, *Survey of Current Business*, Apr., 1958, and Apr., 1964.
12. T. PAUL SCHULTZ, *The Distribution of Personal Income* (Joint Economic Committee, Dec., 1964).
13. ———, "Secular Trends and Cyclical Behavior of Income Distribution in the U.S., 1944-1965," *Six Papers on the Size Distribution of Income and Wealth*, Studies in Income and Wealth, v. 33 (N.B.E.R., 1969), pp. 75-106.
14. ROBERT R. SCHUTZ, "On the Measurement of Income Inequality," *A.E.R.*, Mar., 1951, pp. 107-22.
15. LEE SOLTOW, "The Share of Lower Income Groups in Income," *Rev. of Econ. and Statis.*, Nov., 1965, pp. 429-33.
16. Survey Research Center, *Survey of Consumer Finances* (Univ. of Michigan, 1960-67, and *Fed. Res. Bul.*, Vols. 35, 36, 45).

ANALYZING THE AMERICAN INCOME DISTRIBUTION*

By LESTER C. THUROW

Massachusetts Institute of Technology

What causes changes in the American income distribution? Typically existing analysis focuses on explaining changes in the poverty percentage or differences in median incomes across racial groups.¹ Such studies provide limited information about the factors which alter the distribution of income. Factors may have one impact on median incomes and very different impacts on the incomes of the rich or poor. Thus, summary measures of income distribution, such as medians, can be misleading.² Finally, the factors that cause changes in the distribution of income are themselves distributions. The distribution of education and training affects the distribution of income. Thus, to adequately study the American distribution of income, it is necessary to develop methods of explaining the distribution of income in terms of the distributions of causal factors which influence it.

To investigate the American income distribution, it is necessary to develop some technique for describing and manipulating entire distributions. One answer is to fit analytic distributions to the observed distribution of income and its causal factors. An attempt can then be made to relate the parameters of these distributions to each other. The income distribution parameters are explained by the education parameters, training parameters, and whatever other factor may be important in explaining the distribution of income.

There are two broad states to such a study. The first is to analyze the impact of macroeconomic factors on the distribution of income. This is an area pioneered by the work of Charles Metcalf.³

The second state is to determine the impact of microeconomic distributions on the distribution of income. This paper is part of a broader study focusing primarily on the second stage of this process, but it is first necessary to analyze the different impacts macroeconomic factors have on the income distributions of blacks and whites. This paper focuses on the macroeconomic factors over time and across states. At the present stage of the project, distributions have not yet been fit to the microeconomic explanatory factors.

I. The Beta Distribution

A number of analytic distributions can be used to describe changes in the distributions of income for blacks and whites,⁴ but the beta distribution seems the most flexible.⁵ It fits the observed income distributions well (see below) and has only two parameters. These two parameters are p and σ (see equation (1)). By placing incomes on a scale between 0 and 1, the proportion of the population, p , who have some particular income is easily calculated (see equation (2)).

COMPLETE BETA FUNCTION

$$(1) \quad \beta(p, \sigma) = \int_0^1 t^{p-1}(1-t)^{\sigma-1} dt$$

$$p > 0$$

$$\sigma > 0$$

$$(2) \quad f\beta(p | p, v) = \frac{1}{\beta(p, \sigma)} p^{p-1}(1-p)^{\sigma-1}$$

where

$$0 \leq p \leq 1$$

$$\sigma = v - p$$

$$p, \sigma > 0$$

* Part of the work on this paper was financed under a National Science Foundation grant, NSF-GS-2811.

¹ This is true of my own work as well as that of others.

² For example, existing analysis of the impact of education on incomes looks at the difference in mean or median incomes between education cells. The analysis explicitly assumes that there is no income variance within education cells; yet the actual income variance within education cells is often greater than the variance between cells. Often medians are not statistically significantly different. Both regression techniques and tabular techniques for isolating the returns to education have this problem. If individual observations are used, the coefficient of determination in regression analysis provides some indication of the magnitude of the problem, but the regression coefficients are based on mean values.

³ Charles Metcalf, "The Size Distribution of Personal

Income in an Econometric Model of the U. S." (unpublished thesis, M.I.T.).

⁴ Metcalf uses what he calls the displaced log normal distribution. This distribution has three parameters with an infinite upper tail.

⁵ Eventually I wish to use the beta distribution to describe both independent and dependent variables. Using a beta distribution is not meant to imply that God is a beta generating function. For a more detailed discussion of the beta function see: M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics* (Charles Griffin and Co., Ltd., London, 1963), p. 150; or John W. Pratt, Howard Raiffa, and Robert Schlaifer, *Introduction to Statistical Decision Theory* (McGraw-Hill Book Co., 1965), p. 91.

The impact of changes in ρ and σ on the distribution of income is easily determined. Increases in σ (see Appendix, Chart I) result in higher median incomes and smaller relative income differences between the top and bottom of the income distribution. In the ranges under consideration in this paper a 1 percent increase in σ (holding ρ constant) results in an approximately 0.8 percent increase in the median income and a 0.4 percent reduction in the relative income gap between the 25th and 75th percentile of the income distribution. Increases in ρ (see Appendix, Chart II) result in lower median incomes and larger relative income differences in the income distribution. A 1 percent increase in ρ results in a 0.4 percent reduction in the median incomes and a 1 percent increase in the relative income gap between the 25th and 75th percentile of the income distribution. When both ρ and σ rise in the same proportion (see Appendix, Chart III) median incomes remain constant but relative income differences increase. The dispersing impact of ρ dominates the concentrating impact of σ . More precisely, a 1 percent increase in both ρ and σ results in a 0.2 percent increase in the income gap between the 25th and 75th percentiles of the income distribution.

II. Estimating ρ and σ

The beta distribution was fit to U. S. Bureau of Census constant dollar (1959) income distribution statistics for households (families and unrelated individuals) for every year from 1949 to 1966.⁶ Any number of observations could have been used in the estimating process, but ten were actually used. These ten observations were the proportion of the population who had incomes less than 0.05 through 0.95 of the income scale. Since the beta distribution is a finite distribution, it has a maximum income implicit in its estimation. The maximum income can be adjusted, but it was set at \$15,000 in this work. Thus everyone who has an income over \$15,000 is given an income of \$15,000. No one is completely ignored. This is not a serious constraint, since less than 5 percent of the population have incomes over \$15,000 (1959 dollars) in 1966, and their first \$15,000 is being considered. A \$25,000 income limit works just as well, but seemed inappropriate over the period under consideration. The proper income limit depends upon the focus of the anal-

ysis. The larger the income limit the more weight is put upon the upper tail of the income distribution in the estimating process.⁷ Since this analysis is not focused on the factors which influence the incomes of the wealthy, a relatively low income limit was chosen.

Between 1949 and 1966 the median incomes of white households rose from \$3,796 to \$6,084 (1959 dollars) and from \$1,885 to \$3,429 for black households. Measured in terms of Gini coefficients⁸ or relative incomes there have been almost no changes in the distribution of either white or black incomes (see Appendix, Chart IV).

The beta distribution fit the actual income distributions well. The coefficient of determination (R^2) improves over the period, but averages .96 for whites and .92 for blacks. (The actual parameters are available upon request from the author.) Rho and sigma rise for both whites and blacks. Median incomes have grown and the dispersion of income has remained relatively constant over the period under consideration, but this was a product of two offsetting forces. Increases in ρ lead to a greater income dispersion and increases in σ lead to less income dispersion. While the distribution impacts of ρ and σ offset each other, the σ dominated ρ in terms of growth in absolute income.

The ρ 's and σ 's differ substantially between blacks and whites. Rho is relatively higher for blacks while sigma is lower. This is a reflection of the fact that the income distribution for blacks is much more disperse than that for whites. The same differences can be seen in Gini coefficients. The Gini coefficient for blacks is substantially higher than it is for whites.

Growth of the two parameters is also substantially different. The white ρ rose from .666 to 1.044 while the black ρ rose from .930 to 1.104. The white σ rose from .258 to .955 while the black σ rose from .160 to .514. Thus by the end of the period, the ρ 's are rather similar, but the black σ is only half as large as the white σ . Thus, most of the differences in income level and dispersion is produced by the σ parameter.

The beta distribution works almost as well across states in 1960 as it does over time. Among whites all of the R^2 are above .95 and most are above .98. The results are not quite as good among nonwhites. Of 51 observations the R^2 is above .90 in 33 and below .80 in only 2. There is a sig-

⁶ The income data are from: U. S. Bureau of the Census, *Trends in the Income of Families and Persons in the United States, 1947 to 1960* (USGPO), Table 11; U. S. Bureau of the Census, *Current Population Reports, Consumer Income*, Series P-60, various issues.

⁷ As the maximum income limit increases, the number of upper income observations used in estimating ρ and σ increases.

⁸ The Gini coefficient is the area between the diagonal and the Lorenz curve divided by the area under the diagonal.

nificant amount of variation in both ρ and σ for whites and nonwhites across states.⁹

III. Factors Producing Changes in ρ and σ Over Time

In this analysis the impact of macroeconomic factors is separated into five parts. There is a real growth component (measured in terms of constant dollar GNP per employee, GNP/E); an inflation component (measured in terms of the implicit price deflator for GNP , I); a factor shares component (measured in terms of the share going to personal income, PI/GNP); a utilization component (measured in terms of the proportion of the labor force employed, E/LF); a transfer payment component (measured in terms of the transfer payments per household, TP/H); and a government expenditures component (measured in terms of government purchases of goods and services per household, GP/H).¹⁰ The model is given in equation (3).¹¹

$$(3) \quad \rho \text{ or } \sigma = A \left(\frac{GNP}{E} \right)^{b_0} (I)^{b_1} \left(\frac{PI}{GNP} \right)^{b_2} \left(\frac{E}{LF} \right)^{b_3} \cdot \left(\frac{TP}{H} \right)^{b_4} \left(\frac{GP}{H} \right)^{b_5}$$

where

- A = intercept
- GNP = gross national product
- E = total employment
- I = implicit price deflator for GNP
- PI = personal income
- LF = total labor force
- TP = transfer payments
- H = households
- GP = government purchases of goods and services
- $b_0 \dots b_5$ = elasticities

The results of estimating equation (3) are given in Table 1. Whenever variables had t -statistics less than 1, that variable was dropped from the model. The same variable can, of course, be in-

strumental in explaining both ρ and σ .¹² Growth can lead to both a more dispersed and a more concentrated income distribution. To determine the net impact of growth on the income distribution, its impact on the two parameters must be combined. Thus for whites the elasticity of ρ with respect to GNP/E is 1.20 and the elasticity of σ with respect to GNP/E is 2.94. Since the income dispersing effect of ρ is roughly twice as large as the income concentrating effect of σ , the income dispersing effects cancel each other (see above). Growth leads to higher real incomes, but does not have any major impact on the dispersion of income.

The rate of inflation also has two conflicting tendencies for whites. A 1 percent increase in inflation results in a 1.55 percent reduction in ρ and a 0.64 percent reduction in σ . Reductions in ρ lead to a more concentrated income distribution and reduction in σ leads to more dispersed income distribution. Since the ρ effects are roughly twice as powerful as the σ effects in terms of income dispersion, inflation seems to be a powerful force leading to a more equal real distribution of income.

The offsetting impacts of business cycles on white incomes can be seen in the PI/GNP and E/LF terms. In booms employment rises but the proportion of income going to persons falls. Rising employment leads to a lower ρ and more equality, but a falling share for personal income leads to a rising ρ and more inequality. Because of these two offsetting forces the distribution of white income does not change markedly over the course of a business cycle.

Transfer payments present a bit of a problem for whites.¹³ They have an impact on both ρ and σ , but the results indicate that increasing transfer payments leads to a greater dispersion of incomes for whites. Since most government transfer payments do not go to the very rich or the very poor, such a result is not impossible but it is surprising. Evidentially government pensions, veterans benefits, and social security benefits to the middle class dominate welfare payments to the poor for whites.¹⁴

⁹ Cross-sectional state data are from the 1960 U. S. Census. The actual parameters are available from the author.

¹⁰ These data are standard National Income and Products Account data and standard labor force data taken from *Employment and Earnings*, U. S. Dept. of Labor. The employment rate is not the normal unemployment rate since it is total employment divided by the total labor force rather than civilian employment divided by the civilian labor force.

¹¹ The first four terms in this equation yield current dollar percent income per member of the labor force.

$$\left(\frac{GNP}{E} \right) \left(\frac{PI}{GNP} \right) (I) \left(\frac{E}{LF} \right) = \frac{PI}{LF}$$

¹² Since ρ and σ are independent of each other, simultaneous estimation techniques are not necessary.

¹³ In 1968 government transfer payments totaled \$48.6 billion. Direct relief was only \$4.9 billion.

¹⁴ The index is defined as follows:

$$I = \sum_{i=1}^n X_i W_i$$

where

X_i = percentage of the state's labor force in industry i .
 W_i = the ratio of the U. S. median income in industry

TABLE 1
TIME SERIES
DISTRIBUTION ELASTICITIES

	White		Black	
	ρ	σ	ρ	σ
<i>A</i>	3.97 (2.44)	-3.24 (2.21)	-1.47 (.93)	.05 (.50)
<i>GNP/E</i> b_0	1.20 (.60)	2.94 (.43)	-.73 (.18)	
<i>I</i> b_1	-1.55 (.57)	-.64 (.55)	.68 (.23)	
<i>PI/GNP</i> b_2	-2.43 (1.50)			1.56 (1.18)
<i>E/LF</i> b_3	-2.52 (1.94)			6.23 (1.37)
<i>TP/H</i> b_4	.49 (.22)	.39 (.13)	.26 (.07)	.73 (.06)
<i>GP/H</i> b_5				.37 (.08)
<i>R</i> ^a	.96	.99	.95	.99
<i>S</i> _o	.04	.05	.02	.04
d.w.	1.82	1.59	2.13	2.05
d.f.	12	14	14	13

Such is not the case for blacks. Transfer payments influence both ρ and σ , but the impact on σ is three times as large as that on ρ . Consequently transfer payments lead to a more equal distribution of black income. Given the size of central city welfare payments such a result is not surprising.

Government purchases of goods and services lead to higher and more equal black incomes whereas they had no impact on the level or distribution of white incomes. Direct and indirect employment on government projects is one of the major causes of increases in black incomes.

Business cycles have conflicting impacts on the level and distribution of black incomes, but the impacts do not balance out as they did for whites. The employment elasticity (b_4) is four times as large as the share elasticity (b_5). Full employment is a powerful force leading to higher black incomes

and more equal black incomes since blacks suffer more than their share of unemployment and consequently are differentially aided by its elimination.

Inflation, however, leads to more unequal black incomes. Instead of being negatively related to ρ as it is for whites, inflation is positively related to ρ for blacks. Since black incomes are almost entirely wage and salary earnings and since rich blacks are lower middle class by white standards, results may simply indicate that the incomes of those people who are lower middle class by white standards rise relatively in inflationary periods. Thus, if inflation and employment are related to each other as in the Phillips curve, part of the equalizing impact of full employment will be offset by the concomitant inflation, though employment elasticity is almost 9 times as large as the inflation elasticity. Adjusting for the differences in the impact of ρ and σ on the distribution of income still means that there is a powerful equalizing impact from full employment.

General growth has a different impact on black

^a to the general U. S. median income. This index measures the prevalence of high-wage industries in the state.

TABLE 2
CROSS-SECTIONAL DISTRIBUTION PARAMETERS

	<i>a</i>	Percent of Families Living on Farms	Percent of Families with No One in Labor Force	Percent of Family Heads with 0-7 Years of School Complete	Percent of Population 14 and Above Who Worked 50-52 Weeks per Year	Index of Industrial Structure	State and Local Expenditures per Capita	\bar{R}^2	d.b.	S_e
<i>Whites</i>										
ρ	.4414 (.1149)	.4723 (.1133)	1.9053 (.3045)	.3441 (.0985)	.3915 (.2661)		.0351 (.0153)	.99	45	.006
σ	-1.1894 (.2126)			.1636 (.0992)	.6776 (.2476)	1.3042 (.1995)	.0678 (.0145)	.98	46	.006
<i>Non-whites</i>										
ρ	3.0180 (.3178)				-.9257 (.4086)	-1.6691 (.3381)		.97	48	.012
σ	-.2792 (.2036)			-.4363 (.0950)	.3849 (.2371)	.4837 (.1910)	.0248 (.0139)	.94	46	.006

and white incomes. Growth (GNP/E) is negatively related to the black ρ parameter rather than being positively related to both ρ and σ as for whites. For whites, growth leads to higher incomes but did not have much, if any, net impact on the distribution of income. For blacks, growth leads to higher real incomes, but it also leads to more equal incomes. This occurs since black incomes, rich and poor, are dependent on wage and salary earnings. Thus, the growth of capital incomes which prevents growth from equalizing white incomes does not affect black incomes.

Some of these results are based on a loose definition of statistical significance (t statistics greater than 1), but the important results are based on coefficients which are easily significant at the 1 percent level. These include the growth effects, the inflation effects, the transfer payments effects, and the government purchases effect.

IV. Factors Producing Changes in ρ and σ Across States

The cross-sectional model relies on six explanatory variables: the percentage of families living on farms, the percentage of the population 14 and above who work 50-52 weeks per year, an index of the industrial structure, the percentage of families with no one in the labor force, the percentage of family heads with 0 to 7 years of education, and state and local general expenditures per capita (see equation (4)).¹⁵ The first two factors

¹⁵ The independent variables refer to the state as a

represent the level of productivity in a state, the second two represent the labor force participation characteristics of a state, the fifth represents the quality of the labor force, and the final characteristic is designed to reflect the quality of public services in a state.

The model was applied to the 51 state observations using weighted regressions.¹⁶ The results are given in Table 2.

$$(4) \quad \rho \text{ or } \sigma = a + bF + cL + dE + eW \\ + fI + g\left(\frac{E}{C}\right) + u$$

where:

F = percentage of families living on farms

L = percentage of families with no one in the labor force

E = percentage of family heads with less than eight years of school completed

P = percentage of population fourteen years old and above who worked 50-52 weeks per year

I = an index of the industrial structure of the state

whole and not to either the white or nonwhite populations in the state.

¹⁶ A regression using equation (4) needs to be weighted by the population of each state, since the dependent variable and most of the independent variables are in percentage terms. Since a large state provides more of the total number of families living in poverty, it needs to have a larger weight in the regression. The intercept term is a scaling f factor in a weighted regression.

u = error term

E/C = state and local government general expenditures per capita

For whites, increasing the percentage of families living on farms and percentage of families with no one in the labor force leads to lower median incomes and more dispersed incomes. Neither of these two factors seemed to have any impact on nonwhite income distributions. Because of lower black incomes, being on a farm or in a family with no one in the labor force is not the relative handicap it is for whites.

Increasing the percentage of the population working full time and improving the industrial structure leads to higher median incomes and less income dispersion for whites and blacks, but the effects are much stronger for blacks than for whites. A 1 percentage point increase in full-time work and in the industrial structure would raise white incomes approximately 1.4 percent and reduce the income differential between the 25 and 75 percentiles by 0.4 percent. For blacks the corresponding improvements are 1.7 percent and 2.9 percent. Being at the end of the hiring queue, blacks gain relative to whites with tighter labor markets and improvements in industrial job opportunities. The difference between rich and poor blacks depends upon the supply of jobs; the difference between rich and poor whites does not depend on the supply of jobs to the same degree.

Increasing state and local government expenditures per capita raises the incomes of whites without affecting the dispersion of income, but for blacks increasing public services both raises their incomes and reduces the dispersion in incomes. Better public services probably increase the probability of finding jobs for blacks and this reduces the income dispersion among blacks. For richer whites, jobs to do have the same impact.

Education plays its expected role for blacks. The greater the proportion of poorly educated people in a state the lower are median black incomes and the greater is the income dispersion among blacks. Among whites increasing the proportion of poorly educated people does not have a great impact on the median family income but it

does increase the dispersion of incomes substantially. The median white family head is not poorly educated and thus changes in the proportion of poorly educated people over the ranges that are encountered in the 51 states do not affect his income.

The effective instruments for closing the income gaps between blacks and whites are providing full-time jobs and improving the industrial index (productivity levels) of a state.

V. Conclusions

As the models stand there is no effective macroeconomic instrument for equalizing black and white incomes though there are instruments for making them more equal. General growth results in higher incomes for both blacks and whites but does not have much impact on their relative incomes. A 1 percent increase in employment results in a 5 percent increase in median black incomes and a 1 percent increase in median white incomes, but this is an instrument that is inherently limited. Unemployment can only fall to zero. Transfer payments and government purchases lead to more equal incomes, but the coefficients are small enough that enormous expenditure increases would be necessary to equalize incomes using the current structure of transfer payments and government incomes. Regional economic policies which would increase job opportunities in relatively depressed states and which would improve the quality of the existing jobs can also have an impact on black-white income differentials.

If society simply wishes to narrow its income distribution, any factor which lowered ρ or raised σ would be acceptable. Growth would be a proper policy instrument for blacks and inflation would be a proper policy instrument for whites. Realistically, macroeconomic policies can lead to a more equal distribution of income, but vigorous use of these instruments would still leave a widely dispersed income distribution.

The major factor leading to increases in average incomes is, of course, growth for either blacks or whites.

APPENDIX

DISTRIBUTION PARAMETERS AND THE DISTRIBUTION OF INCOME

CHART I

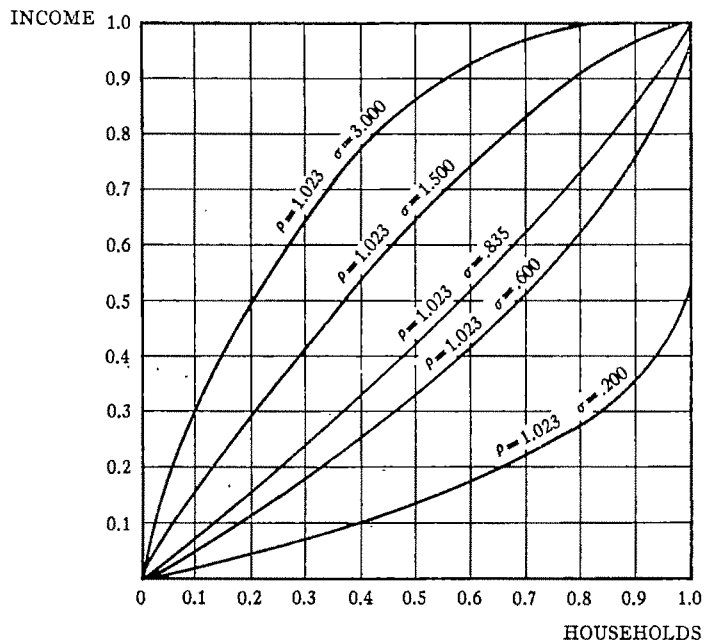


CHART II

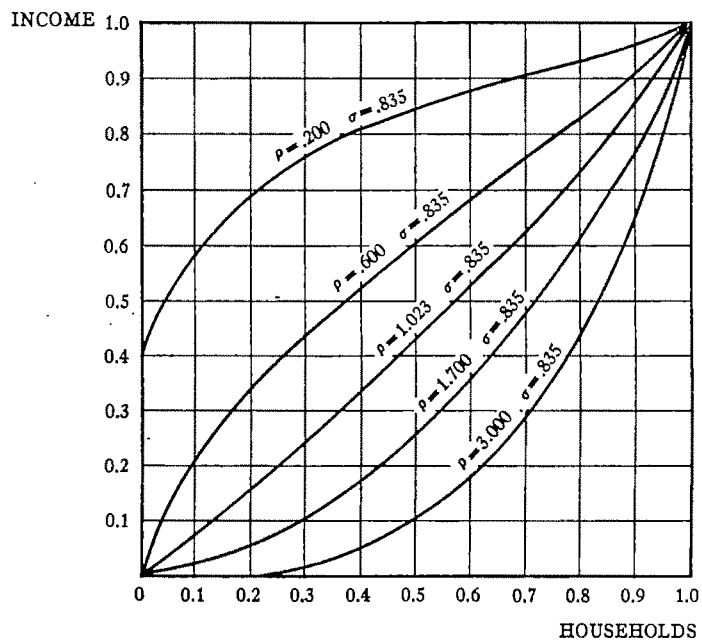
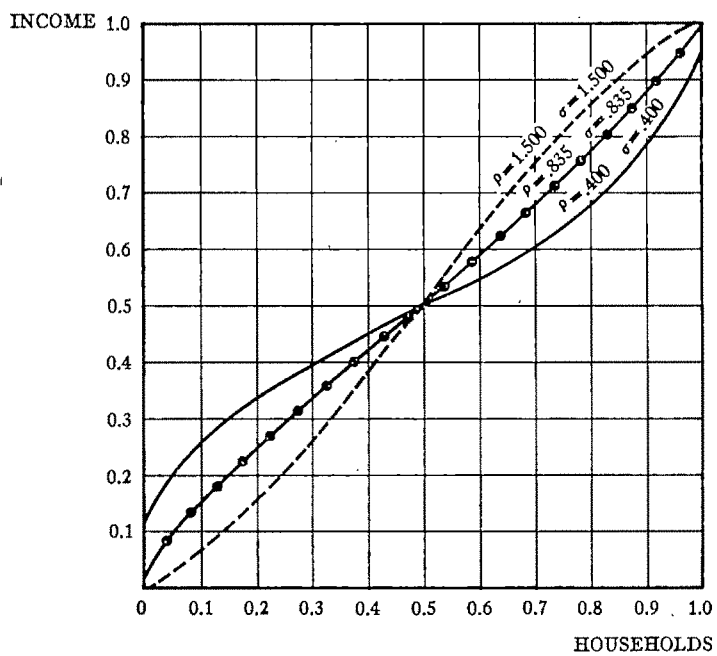


CHART III

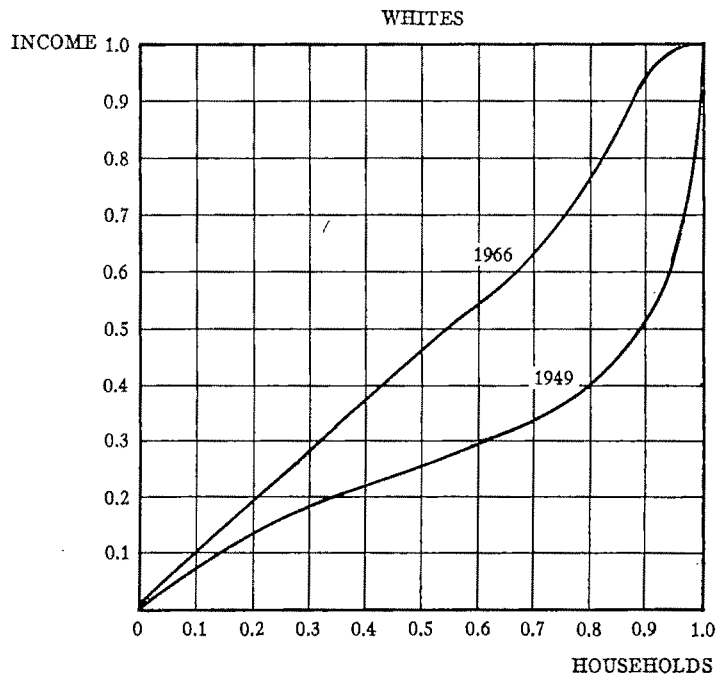
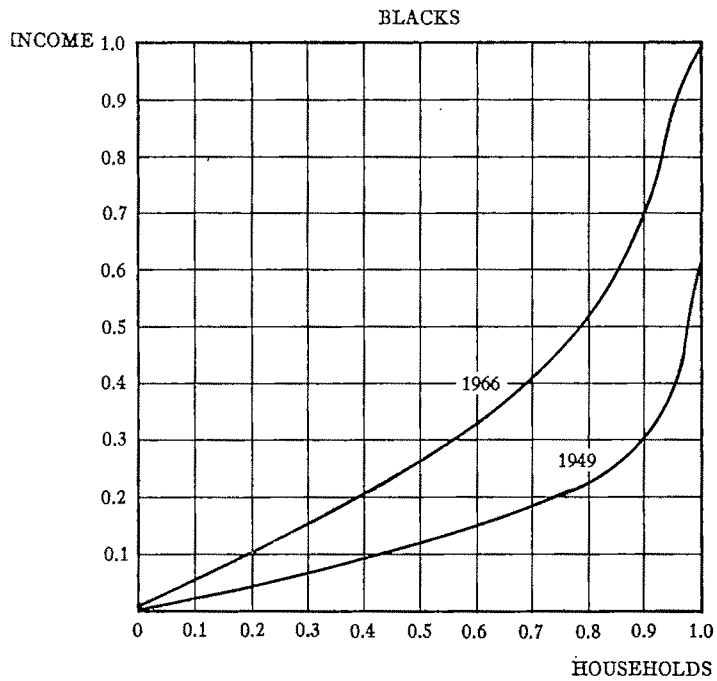


Scale

Households 1.0=100 Percent of Households

Income 1.0=\$15,000

CHART IV
ACTUAL INCOME DISTRIBUTIONS
(Constant 1959 Dollars)



Scale

Households 1.0 = 100 Percent of Households
Income 1.0 = \$15,000

TRANSFER APPROACHES TO DISTRIBUTION POLICY*

By ROBERT J. LAMPMAN
University of Wisconsin

There are two ways to alter the pattern of economic inequality among persons. One is to modify the distribution of factor income by changing the underlying distribution of factors or the prices or employment of those factors. The second is to modify the process by which factor income is redistributed away from its recipients. This paper is about the second way. We ask how the distribution to factor owners is and can be modified as income moves from its market origins to its disposition on goods and services. At the outset, we sketch an accounting framework within which to envision this process whereby "producer incomes" are transformed into "user incomes." In later sections, we review some current proposals for additional transfers to the poor.

The Process of Redistribution

The redistributive process involves receipts and payments of both a money and a nonmoney character. It occurs via private as well as public institutions, including the employer, the government agency, the private insurance carrier, the private philanthropic agency, and the family. Two types of redistributive transaction are involved; namely, transfer and what we may call distributive allocation. Table 1 (see list of items and column 1; ignore columns 2 and 3 for the moment) sketches out the elements of and offers rough estimates of the amounts involved in this process, which may be said to start with subsidy to factor incomes. It continues with public subsidies to consumers, money transfers, and distributive allocations, and with private gifts, transfers, and distributive allocations. Our accounting includes the payments for pure public goods, but, in effect, excludes the benefits of those goods on the grounds that such benefits are indivisible and hence not redistributive. It also excludes those merit-want goods which are produced in the public sector but are purchased on a user-charge or benefits-received basis and hence, like most private purchases, have no important redistributive effect.

* The research reported here was supported by funds granted to the Institute for Research on Poverty at the University of Wisconsin by the Office of Economic Opportunity, pursuant to the provisions of the Economic Opportunity Act of 1964. The author was assisted in the preparation of this paper by Benton P. Gillingham. Earlier writings by the author on this topic include [9] [10] and [11].

As seen in Table 1, the items which intervene between preredistribution income and postredistribution income all have a positive and a negative side. Receipts by one family are canceled out by payments made by others. Hence, items 7, 14, and 18 will be zero in value for the nation as a whole. However, an individual family may have either a positive or negative balance in any one of these items and may gain or lose in the conversion from pre- to postredistribution income. The ranking of families and the overall inequality may be markedly different in the two distributions.

There is interest in the intermediate income concepts shown in the table. For example, one might want to know how the preredistribution income is altered by the public policies reflected in items 5 and 6, and therefore look to item 8. Or, one might want to investigate the effect of private activities as indicated by items 14 and 15.

There are several conceptual issues that surround this accounting scheme. One has to do with the inclusion of nonmoney items. Can we say that the purchase by a public agency of health care for a citizen is not only an allocation but also a "distributive allocation" or "transfer-in-kind"? Does it constitute an addition to "income" for the recipient rather than (or as well as) an investment in his future productivity? My answer to these questions is yes, but I must admit that drawing the line between what is and what is not an in-kind transfer to persons is hard. The line has to do with the distinction between merit wants and social wants, which, as Musgrave puts it, turns on whether the want can be satisfied for one person exclusively. It has to do with Ida C. Merriam's definition of "social welfare expenditures" as those that are directly concerned with income security, and the health, education, and welfare of individuals and families, but exclusive of community-wide utilities and services—such as water and sewer works, urban transportation, or public recreational facilities [14, p. iii].

Another conceptual issue is: should private and social insurance for health care and income maintenance be included? Is the purchase of insurance something that should be accounted for as a simple allocation, like the purchase of postage stamps or automobiles and hence excluded from Table 1? In one view, the benefit is realized

TABLE 1

PUBLIC AND PRIVATE TRANSFERS AND DISTRIBUTIONAL ALLOCATIONS RECEIVED AND PAID BY ALL HOUSEHOLDS AND BY PRETRANSFER POOR HOUSEHOLDS, 1967

	All Households (Billions of Dollars) (1)	Pre transfer Poor Families	
		Percent of Column 1 (2)	Billions of Dollars (3)
1. Preredistribution income (factor income net of 2 and employer financed part of 9).....	644	3	19.3
2. Increase in factor income due to direct subsidy.....	1	10	0.1
3. Reduction in market price due to direct consumer subsidy.....	1	10	0.1
4. Benefits of social welfare expenditures under public programs.....	100	40	40.1
<i>a.</i> Social insurance.....	37	55	16.5
<i>b.</i> Public aid.....	9	93	8.4
<i>c.</i> Veterans.....	7	46	3.2
<i>d.</i> Other welfare services and public housing.....	3	50	1.5
<i>e.</i> Health.....	8	50	4.0
<i>f.</i> Education.....	36	18	6.5
5. Total of 2-4.....	102	39	40.3
6. Taxes, user charges, fees and public prices used to finance 5.....	102	9	9.2
7. Public transfers and distributional allocations net of 6.....	0	—	31.1
8. Income after public transfer and distributional allocation (1+7)...	644	8	50.8
9. Privately insured benefits related to health and income maintenance.	17	5	0.9
10. Direct interfamily gifts.....	10	50	5.0
11. Gifts via philanthropic institutions	3	33	1.0
12. Total of 9-11.....	30	23	6.9
13. Family and employer payments for insurance, gifts by family.....	30	5	1.5
14. Private transfers and distributional allocations less 13.....	0	—	5.4
15. Income after public and private transfer and distributional allocation (1+7+14).....	644	9	55.8
16. Benefits of general government activity.....	100	8-9*	8-9*
17. Taxes to pay for 16.....	100	7	7
18. Benefits of 16 less taxes of 17....	0	—	1-2
19. Postredistribution income (1+7+14+18).....	644	9	56.8-57.8
20. Allocations other than 4 or 16 through government having no redistributive impact†	—	—	—
21. Allocations through private sector other than 12 having no redistributive impact.....	—	—	—

* Distributed so as to not alter distribution of income. Note pretransfer poor have 8 percent of item 8 and 9 percent of 15.

† Omitted on grounds that these are like private nonredistributional allocations.

TABLE 2

DISTRIBUTION OF FAMILIES AND INCOME, BEFORE AND AFTER FISCAL INCIDENCE OF FEDERAL, STATE, AND LOCAL GOVERNMENT TAXES AND EXPENDITURES, BY MONEY INCOME, 1960

Item	Family Money Income Brackets							Total
	Under \$2,000	\$2,000-2,999	\$3,000-3,999	\$4,000-4,999	\$5,000-7,499	\$7,500-9,999	\$10,000 and Over	
				(percent)				
1. Families.....	14	9	9	11	28	15	14	100
2. "Broad income".....	2	3	4	7	27	19	39	100
3. "Adjusted broad income".	3	5	5	7	26	20	33	100

SOURCE: W. Irwin Gillespie [5]. Line 1 from Table 11; lines 2 and 3 from Table 13.

at the time pure insurance (disregarding cash values) is purchased. One buys protection against risk and gets his money's worth even though he never has a claim. Transfer or subsidy might be said to arise only if premiums are not correctly adjusted to variations in risk or to size of potential benefits. However, insurance is distinguishable from family saving, which we include here only if it is transferred from one family to another by gift. The insurance intermediary, whether it be a private company or a government agency, does something a family cannot do in pooling risk and thereby converting factor income into payments that respond to stated contingen-

cies rather than to accumulated contributions. Moreover, insurance benefits often flow to persons quite remote from the purchasers, who may be employers or general revenue taxpayers. For these reasons, insurance benefits and contributions paid in the year are included in the accounting of distributive allocations of that year.

Benefits and Losses from Redistribution

Having resolved the issue of what is being redistributed and how, we turn now to the matter of who receives how much benefit and who loses how much, from redistribution. One may rank families by total monthly income and measure how the share of income received by income-bracket groups changes as we move from one concept of income to another. Thus, Gillespie [5, see Table 2] moves from factor income (roughly the same as item 1 in Table 1) to a post-redistributional income which takes no account of private transfers and distributional allocations (this is approximately the same as the sum of items 1, 7, and 18). Note that he finds that the share going to those with money incomes under \$3,000 is converted from 5 percent to 8 percent in the redistributive process.

A similar pattern of change is documented by Morgan, David, Brazer and Cohen [15, see Table 3], who reranked families each time they changed income concept. They show that the degree of inequality is reduced one-fifth of the way to zero by moving from gross factor income (about the same as item 1 in Table 1) to gross disposable income (similar to item 19). They also portray a significant shift of another one-tenth of the way to zero inequality, which is associated with a reranking of families by welfare-ratio; that is, adjusting for family size. A welfare-ratio of unity means that a family is at the poverty

TABLE 3

LORENZ COEFFICIENTS OF INEQUALITY FOR VARIOUS UNITS OF ANALYSIS AND MEASURES OF INCOME, 1959

	Families	Adult Units
Gross factor income.....	.419	
Less: Imputed rent of homeowners		
Less: Home production		
Plus: Regular money transfers		
= Money income.....	.385	
Less: Federal income taxes		
= Disposable money income.....	.355	
Plus: Imputed rent of homeowners		
Plus: Home production		
Plus: Nonmoney and irregular transfers including food and housing provided by relatives		
= Gross disposable income.....	.346	
Divided by budget standard		
= Welfare ratio.....	.309	.346

SOURCE: Morgan, David, Brazer, and Cohen [15].
Derived from Table 20-2, p. 315.

TABLE 4
SPECIFIC TRANSFERS AS PERCENTAGE OF GROSS DISPOSABLE INCOME OF ADULT
UNITS AND FAMILIES CLASSIFIED BY WELFARE RATIO, 1969

Transfer Items	Welfare Ratio					
	.0-0.8	0.9-1.2	1.3-1.6	1.7-2.2	2.3 and Over	All Units
<i>^AAdult units</i>						
Nonfamily transfers.....	26	12	9	6	5	8
Income tax.....	-1	-5	-9	-11	-19	-12
Nonfamily contributions..	-4	-4	-4	-4	-7	-5
<i>^BFamilies</i>						
Property tax.....	-2	-2	-2	-2	-2	-2
Public school benefits....	13	7	4	3	1	4
Net transfer.....	26	7	-2	-7	-18	-6

SOURCE: Morgan *et al.* [15]. Panel A derived from Table 16-23. Panel B derived from Tables 19-5 and 19-13.

line. Similarly, they show that the inequality we measure is partly a function of whether we count families "doubled-up" in one household as one or two families. Apparently relatives tend to group together in such a way as to reduce inequality below what we would observe if each adult unit were separately accounted for. It is, of course, a key matter in the study of income redistribution to know how separate factor incomes are combined into a family-unit-income and how persons who do not receive factor income relate themselves to others who do.

Table 4 shows the finding of Morgan and his colleagues that what they call "net transfer," which takes account of only some of the items listed in Table 1, namely, nonfamily transfers plus public school benefits, less income and property taxes (note that they find the latter to be nonredistributive) and nonfamily contributions, amounted in 1959 to 26 percent of the gross disposable income of families below the poverty line.

By making use of the findings presented in the Morgan book and in the studies by Gillespie, Orshansky, McClung, Lurie, Bridges and others, and by reference to some preliminary findings from the Survey of Economic Opportunity, we are able to put together rough estimates of the amount of redistribution to the pretransfer poor¹ done in 1967. These estimates are reported in the right-hand columns of Table 1. They show that the

¹ This means those who have money incomes below the Social Security Administration poverty guidelines before the receipt of money transfers.

whole array of public and private givings and takings raised the share of pretransfer poor families from 3 percent of preredistribution income to 9 percent of postredistribution income. (See items 1 and 19.) This was accomplished by public transfers (positive and negative) and distributional allocations, which raised the share from 3 to 8 percent (item 8); by private transfers and allocations, which raised it to 9 percent (item 15); and by taxes for pure public goods, which raised it by less than 1 percent, so the share remains at about 9 percent (item 19). This gain in share was accomplished by an offsetting decline in share of income on the part of the pretransfer nonpoor families.

Instead of showing how income distribution is modified by redistributive institutions with respect to those in income brackets or welfare-ratio groupings (of which poor-nonpoor is a variant) one could show income redistribution with respect to age, sex, color, region, occupation, family size, home ownership status, health status, labor force status, educational level, or other characteristics of income recipients. Another method of presentation is to show how a particular item, say the income tax, alters distribution among successive groupings of the population. This is the method advocated recently by the Treasury Department [1].

Table 1 suggests the range of approaches for redistributing income to any of the above listed groupings. These include subsidies at the factor income level and in consumer markets, public and private transfers and distributional allocations,

TABLE 5

SOCIAL WELFARE EXPENDITURES UNDER PUBLIC PROGRAMS (TOTAL, AND CASH BENEFITS UNDER PUBLIC INCOME MAINTENANCE PROGRAMS), FOR SELECTED YEARS, 1940-68

Cash Benefits under Public Income Maintenance Programs (In Billions of Dollars)							
Year	Total (In Billions of Dollars)	Total	Retirement, Disability, and Survivors	Unemploy- ment	Temporary Disability	Workmen's Compensation	Public Assistance
1940	8.8	4.2	0.8	0.5	—	0.2	1.0
1950	23.5	8.7	4.3	1.5	0.1	0.4	2.4
1960	52.3	25.9	8.2	3.0	0.4	0.9	3.3
1963	66.8	32.4	24.1	2.9	0.4	1.1	3.6
1967	99.7	42.6	33.4	2.4	0.5	1.4	4.9
1968	112.0	48.2	38.0	2.4	0.5	1.6	5.7
1969	126.8						

SOURCES: Column 1, Merriam and Skolnik [14]. For 1969, Research and Statistics Note, Office of Research and Statistics, SSA, Nov. 12, 1969. All other columns, *Social Security Bul.*, Sept., 1969, Table M-1, p. 33.

and, on the other side of the ledger, taxes, insurance premiums, and gifts.

It is important to know that most of these methods have been increasing in quantitative importance over the years. (Solomon Fabricant discusses reasons for this trend in [4].) Social welfare expenditures under public programs (item 4 in Table 1) have been rising 10 percent or more per year and now amount to 14 percent of gross national product. The increase in the national redistributional effort as indicated by this rise, not having a maximum impact on the number of people counted in poverty nor on the size of the poverty income gap, since many of the greatest rises have been in such nonmoney items as schooling and health care.

However, the cash benefits in this series (see the right-hand section of Table 5) have been rising in step with the total and added up to almost \$50 billion in 1968. Since the latter are the only public expenditures, other than "direct subsidies to increase factor incomes" which immediately enter into total money income, and since the poverty line is stated in terms of preincome tax total money income adjusted for family size, they would seem to have unique relevance out of all the items listed in Table 1 to the question of poverty reduction. It does seem ironic that cash benefits went up from under \$25 billion in 1959 to \$43 billion in 1967, while the poverty-income gap fell only from \$13.7 billion to \$9.7 billion.

Thirty-two percent of all households received a cash transfer in 1965; yet only about 8 percent of

all households were kept out of or taken out of poverty status thereby. See Table 6. We do not have a good series on the number of pretransfer poor, but it does appear that transfers have been taking gradually increasing numbers of households to post-transfer incomes above the poverty line. We found, using 1961 Bureau of Labor Statistics' Consumer Expenditure Survey data, that transfers took 4.7 million households out of poverty in that year [9]. This is the same number counted by Miss Orshansky (see Table 6) for government transfers only for 1965. Irene Lurie found that transfers took 6.1 million households out of poverty in 1966 [12], and we confirm that

TABLE 6

PERCENT OF HOUSEHOLDS RECEIVING SELECTED TRANSFERS AND NUMBER OF HOUSEHOLDS WHO WOULD BE COUNTED POOR BUT FOR TRANSFER, 1965

Transfer Item	Percent of Families Receiving	Number (in millions of families) Who Would Be Counted Poor But for Transfer Payments, Out of a Total of 60.4 Million Households
Any payment...	32.3	4.7
Social security..	21.5	3.6
Public assistance	4.8	.4
Other.....	11.2	1.1

SOURCE: Mollie Orshansky [16, pp. 26-30].

TABLE 7

DISTRIBUTION OF PRETRANSFER INCOME, MONEY TRANSFER BENEFITS AND POST-TRANSFER MONEY INCOME, BY HOUSEHOLD POVERTY STATUS, 1967*

	Total Population (Billions of Dollars)	Pretransfer Poor		Post-transfer Poor (Billions of Dollars)
		Percent of Total	Billions of Dollars	
Pretransfer money income.....	520	3	15.6	7.8
2. Increase in factor income due to subsidy.....	1.0	10	0.1	—
4. Benefits under social welfare expenditure programs (money only).	42.6	57	24.3	8.1
a) Social insurance.....	33.2	53	17.7	4.4
b) Public aid.....	4.9	93	4.5	3.0
c) Veterans.....	4.5	46	2.1	0.7
9. Private insurance (money only)...	10.0	5	0.5	0.2
10. Direct interfamily gifts (money only).....	6.0	33	2.0	1.0
11. Gifts via philanthropic institutions (money only).....	0.5	33	0.1	0.1
Subtotal of 2, 4, 9, 10, 11.....	59.1	46	27.0	9.4
Post-transfer money income (pre-transfer money income plus subtotal above).....	579.1	7	42.6	17.2
Distribution of households (percent)...	100	25		15

* Item numbers correspond to items in Table 1.

number by an independent computation of Survey of Economic Opportunity data.

The increases in money transfers have been largely directed at the pretransfer nonpoor. Of course, the number of people counted as poor has declined dramatically—from 23 percent of the population in 1959 to 12 percent in 1968—but that is apparently due primarily to rising wage rates and improved employment opportunities for low-income people. The unemployment rate fell from a postwar high of 6.8 percent in 1959 to below 4 percent in 1966 and has stayed there since that time.

Nothing said above is meant to minimize the importance of money transfer in reducing poverty. Money transfers, public and private, added substantially to the incomes of the poor in 1967. (See Table 7.) We estimate that about half of the total of \$59.1 billion of such transfers went to the pretransfer poor. They cut the pretransfer poverty-income gap almost in half,² and provided over half of the income of the post-transfer poor.

However, money transfers tend to do more for

² Note that no accounting is made for the possibility that transfers may have induced a reduction in the pretransfer income of some households.

the better-off poor than for the poorest poor. They also tend to favor small families as opposed to larger families among the poor.³ This insight into the bias of the existing set of transfer payments has led many people to advocate new kinds of transfers aimed at children in intact families or "the working poor."

Discussion of New Transfer Plans

The question is persistent, why, in a rich country, having a large and rapidly growing redistributive system, can we not make some adjustment that will simply eliminate the poverty that remains? The dimensions of the problem seem small—only about 23 million people are poor and their poverty-income gap of under \$10 billion is only 1 percent of gross national product. The most direct way to eliminate poverty would be to introduce a negative income tax with guarantees at the poverty lines. If all other income were subject to a special offset tax of 50 percent up to break-even points equal to twice the poverty lines

³ To be fair about this we need to mention that public school benefits and the income tax exemption system modify this bias in money transfers [9, pp. 143 ff.].

for each family size, this would cost \$27 billion in benefits and income tax forgiveness over and above what we are now spending on transfers [17, p. 18]. These net benefits would go to some 88 million people, leaving the upper 120 million to pay the \$27 billion on top of the taxes they are now paying.⁴ Worry about subjecting working people to a 50 percent marginal tax rate—and note that we are talking not about a few categorical poor but about 40 percent of the population—leads some to advocate lowering the offset tax rate to, say, $33\frac{1}{3}$ percent, thereby raising the break-even points to three times the poverty lines, and placing the whole tax load, which would then be expanded to cover about \$50 billion of net new transfers to the lower two-thirds of the population, on the upper one-third. That amount would require a near doubling of the money-transfer now being done by all public and private sources. In any event, what may look at the outset like an easy problem takes on greater scope as one surveys the alternatives. Certainly, it is a major disservice to rational discourse to suggest, as many have done, that the United States could eliminate poverty if we were only willing to transfer an additional \$10 billion to the poor. There is no way to get that \$10 billion into the hands of the poor without spending far more than that.

Realization of this—and some sense of the gradualness of change—has prompted many to offer less radical departures. Perhaps the least radical would modify existing programs. The transfer program now paying out the most cash to the poor is Old Age, Survivors, and Disability Insurance. This could be—and likely will be—expanded, but each extra dollar in benefits tends to yield only 10 to 20 cents for the poor; these benefits do not reach the noncategorical poor; and the payroll tax puts a heavy burden on the working poor. The second largest source of public transfer funds for the poor is the categorical public assistance programs dominated by Aid to Families with Dependent Children (AFDC). We could improve the status of many of the categorical poor by setting a federal floor under the benefits, now determined by the several states. This would cost relatively little, but it would not reach many of the poor—most of the poor are not in the categories—and would exacerbate the inequity between the working poor at low earned in-

comes and the categorical poor at relatively high benefit levels. The only way, via transfers, to remedy that inequity is to drop the age-old principle that receipt of transfers and employability must be mutually exclusive. And one way to broach the contrary principle is to pay benefits to all children. Since most of the nonwelfare poor are in families with children, a child allowance, which is a common type of transfer in other nations, would seem to have time and place utility for the United States.

There are numerous varieties of child allowance plans, but like all transfers, they take from some and give to others, most obviously transferring income from households without children to those with children. They also tend to alter the distribution of income among families with and without children. Benefits can be conditioned in various ways and can be financed, at least in part, by offsetting reductions in existing ways of changing disposable income such as cash transfers and the exemptions for children in the income tax. In order to see how plans differ from one another it is useful to do, as Dorothy S. Projector [18] has done, a calculation of how disposable income would change for the average family in each income bracket. Table 8 shows some of her calculations for 1967 for four plans; namely, two suggested by Vadakin [20], one by Brazer [3], and the Family Assistance Plan (FAP) recommended this fall by President Nixon and introduced by Congressman John Byrnes as H.R. 14173. We have added parallel calculations for 1966 for the plan recommended in November by the Presidential Commission on Income Maintenance Programs (the Heineman Commission) [17], even though their plan would pay benefits to unrelated individuals and families with and without children and hence is not a child allowance. We would like to include a wage subsidy in this comparison but do not know of a carefully spelled out plan of that type. (See Kesselman [8] for a discussion of the issues.)

The several plans, briefly sketched, are as follows. To achieve comparability, we assume, following Projector, that each plan is to be financed, to the extent new tax revenue is needed, by a surtax on personal and corporate income.

Vadakin 1: \$120 per year allowance per child; retain child exemption in the income tax; finance by making allowance taxable and adding a 7.4 percent surtax on personal and corporate income.

Vadakin 2: \$120 per year allowance per child; finance by eliminating child exemption and making allowance taxable. (No surtax needed.)

Brazer: \$1,400 per year allowance for first child, \$900 for second, \$600 for third, \$400 for each

⁴ It is interesting to recall Harry G. Johnson's 1964 comments on a transfer of this amount. He said that it "... may well be politically unacceptable; but it is really small potatoes as war finance goes, if war on poverty is really what has been declared" [7, p. 545].

TABLE 8

AVERAGE AMOUNT OF CHANGE IN DISPOSABLE INCOME FOR FAMILIES AND UNRELATED INDIVIDUALS, BY TOTAL MONEY INCOME, FOUR CHILDREN'S BENEFIT PLANS, 1967, AND HEINEMAN PLAN, 1966

Total Money Income	Vadakin 1	Vadakin 2	Brazer	Family Assistance Plan	Heineman
Under \$3,000.....	\$ 49	\$ 47	\$ 260	\$ 127	\$ 417
\$3,000-4,999.....	78	30	345	75	184
\$5,000-6,999.....	78	2	139	-9	-7
\$7,000-9,999.....	63	-15	-132	-29	-151
\$10,000-14,999.....	3	-30	-211	-49	-246
\$15,000-24,999.....	-159	-59	-346	-102	-457
\$25,000 and over...	-845	-147	-949	-347	-1517

SOURCES: Columns 1-4, Dorothy S. Projector [18], Table 5. Column 5, Nelson D. McClung [13].

added child; finance by eliminating child exemptions, taxing adjusted gross income by a special child allowance tax at marginal rates around 33 percent but varying both by income and family size, reducing federal contribution to AFDC, and adding a surtax of about 6 percent on personal and corporate income.

Family Assistance Plan: Benefits restricted to families with children; \$500 per year for first two persons, \$300 for each additional person; finance by taxing other income by a special offset tax at a zero rate on first \$720 of earnings and at a 50 percent rate beyond that, eliminating federal contribution to AFDC, adding a surtax of about 2.5 percent on personal and corporate income.

Heineman: Benefits not restricted to families with children; \$750 per year for each of first two adults, \$450 for each other person; finance by taxing other income by a special offset tax at a 50 percent rate, eliminating federal contribution to food stamps and to all categorical assistance programs, adding a surtax of about 12 percent on personal and corporate income. (The latter surtax is equivalent to a surtax of 18 percent on personal income only. These surtax rates are for 1966 and would be lower for 1967.)

Table 8 demonstrates the similarity in basic design of all child allowance and negative income tax plans in changing disposable income by income bracket. The break-even points differ, ranging from around \$15,000 in the Vadakin 1 plan to \$5,000 in the Family Assistance Plan. The amount of gain and the distribution of that gain, as well as the amount and distribution of loss, vary among the several plans. Further insight into the variations of the plans is offered by Table 9, which shows the total increases in disposable in-

come occasioned by the introduction of the plans. It should be noted that these estimates take no account of possible reductions in work effort nor of changes in family size or composition which might result from the plan. These increases in income, which are matched by decreases above the line, vary from a low of \$1.1 billion for Vadakin 2 to a high of \$8.6 billion for the Heineman plan. The net benefit to the poor also is different among the plans. The Family Assistance Plan does the most for the poor per dollar transferred and in that sense may be said to be the most intensively antipoverty plan. A proponent of child allowances has commented on this point as follows: "It will be said that a child allowance wastes money on children who are not poor. . . . A child allowance designed carefully in relation to the income tax system would waste little money. In any event, that money is well wasted that purchases a sense of its rightness. . . . [Moreover] because it is not related to income it quite avoids interfering with the incentive to work" [19]. (Note, however, the high guarantee and high marginal tax rates in the Brazer plan.) None of these plans would confine its benefits to the 23 million persons who are poor. FAP would add 14 million persons to those in benefit status, some of whom already have incomes above the poverty lines. The Heineman plan would reach 36.8 million persons in 1971, almost half of whom would not be poor in the absence of the benefits.

Deciding on how to rank these plans in terms of desirability may well turn, for each citizen, on a complex set of considerations. It may lend perspective to relate the amount that would be transferred under these plans to the amount now transferred. According to Table 7, the total amount

TABLE 9
COMPARISON OF FOUR CHILDREN'S BENEFIT PLANS, BASED ON 1967 INCOME
AND THE HEINEMAN PLAN, BASED ON 1966 INCOME

Characteristic	Vadakin 1	Vadakin 2	Brazer	Family Assis- tance Plan	1966 Heine- man
1. Amount "transferred"* (billions of dollars).....	2.5	1.1	7.5	2.5	8.6
2. Income level above which average change in disposable income is negative (see Ta- ble 8).....	15,000	7,000	8,000	5,000	6,000
3. Effect on those without children.....					
a) Receive benefit.....	No	No	No	No	Yes
b) Pay tax.....	Yes	No	Yes	Yes	Yes
4. Net benefit to households with under \$3,000 income (billions of dollars).....	0.6	0.6	3.6	1.8	5.6
5. Increase in share (was 4.1 percent) of income going to to lowest fifth of households (percent).....	0.1	0.1	0.4	0.3	1.2
6. Marginal tax rates below break-even income (per- cent).....	†	†	30-40	0-50	50

* In this case, "transfer" means change in disposable income. The total amount of such change below the levels shown in line 2 is matched by an offsetting amount above the line.

† Surtax only.

SOURCE: On the Heineman Plan [6] [13] and [17]. On the other plans, calculated from data in Projector [18].

of money transfers in 1967 was \$59.1 billion. So, the largest amount listed here, the Heineman plan's cost for 1966 of \$8.6 billion, is about one-seventh of that total. (The amount needed to finance that plan in 1971 is estimated to be only \$6 billion.) A similar consideration is suggested by the question: how much change in the share of income going to the lowest fifth of households would follow from each plan? The most redistributive of the plans by this measure is the Heineman plan, which would change that share from 4.1 to 5.3 percent; i.e., a 30 percent increase. (See Table 9, line 5.)

There are, then, certain broad issues to be weighed. How much total transfer, how to divide transfers between cash and in-kind, how much emphasis on children, how intensely to concentrate on poverty reduction, how high a marginal tax rate, what existing transfers should be reduced to help finance any new benefit? But even after those issues are resolved, there are numerous somewhat more technical issues to be settled. A simple family allowance plan, such as Vadakin

1 or 2, does not have to contend with some of these issues, but the other three plans discussed do. We have space here only to list some of them: (1) Should the plan have a work test associated with it, as does FAP? Should the work test apply to all adults? Should the penalty for failing to work less than full time be severe? (2) Should the income subject to the special offset tax be defined broadly (as in FAP) or narrowly (as in the Brazer plan)? Should social security benefits be included and taxed? Should work expenses and child care expenses be deductible? (3) Should the family be defined so as to leave choice as to what persons, and hence whose incomes, are to be included in calculating a family's benefits? (4) What income period should be used in determining benefits? Most negative income tax analysts have assumed a year would be the period, but public assistance administrators use a month. (5) How should a new benefit be articulated with existing public assistance programs? The Heineman proposal is silent on this point, but FAP has

a complicated scheme to assure maintenance of effort while discouraging increases of relatively high benefits by the states for those in the dependent children categories. It also enforces conformity with FAP upon the states in defining income and income period, family, and resource and work tests for eligibility. It also sets maximum combined tax rates for those families simultaneously on FAP and a state benefit program. (6) How should the new cash benefit be related to in-kind benefits such as food stamps and medicaid? If food stamps are priced inversely to income, they take on the basic characteristic of a negative tax and hence have a marginal tax rate associated with them. That marginal tax rate could combine with other tax rates to raise the overall tax rate on some families to very high levels. Should the food stamp bonus be calculated after the FAP benefit but before the state supplementary payment, or after both? The Heine-man Commission urges that all food stamps be dropped and that the funds be diverted to financing their more generous negative income tax. (7) How should the plan be administered and by whom? Should it be handled by the states or by the federal government? If the latter, should it be done by the Internal Revenue Service, the Social Security Administration, or a new agency? Brazier nominates the Internal Revenue Service; FAP points to the Social Security Administration.

Those seven questions indicate the complexity of introducing a new type of income-conditioned benefit into the existing system of transfers. Economists, tax lawyers, welfare administrators and other scholars and experts can help to inform the debate now going on concerning President Nixon's Family Assistance Plan and alternatives to it.

Regardless of how that debate comes out, economists have more to do before we fully understand the set of changing institutions by which we can and do modify the preredistribution income, the goals of redistribution (of which poverty reduction is only one), and the consequences, costs, and benefits of such redistribution.

REFERENCES

1. JOSEPH W. BARR, Secretary of the Treasury, Statement Before Joint Economic Committee of the Congress (Jan. 17, 1969). See especially supplementary statement on "Tax Expenditure: Government Expenditures Made Through the Income Tax System."
2. BENJAMIN BRIDGES, JR., "Current Redistributive Effects of Old Age Income Assurance Programs," Old Age Income Assurance, Part II, U.S. Congress Joint Economic Committee (Dec., 1967).
3. HARVEY E. BRAZIER, "The Federal Income Tax and the Poor: Where Do We Go From Here?" *California Law Rev.*, Mar., 1969, pp. 442-49.
4. SOLOMON FABRICANT, "Philanthropy in the American Economy," *Foundation News*, Sept.-Oct., 1969, pp. 173-86.
5. W. IRWIN GILLESPIE, "Effects of Public Expenditures on the Distribution of Income," *Essays in Fiscal Federalism*, Richard A. Musgrave, ed. (Washington, 1965), pp. 122-86.
6. ROBERT HARRIS, "Role of Taxes and Grants in Income Maintenance," mimeographed paper prepared for meeting of the National Tax Association (Oct. 2, 1969).
7. HARRY G. JOHNSON, "Discussion on the Economics of Poverty," *A.E.R.*, May, 1965, pp. 543-45.
8. JONATHAN KESSELMAN, "Labor Supply Effects of Income, Income-Work, and Wage Subsidies," *J. of Human Resources*, Summer, 1969.
9. ROBERT J. LAMPFMAN, "How Much Does the American System of Transfers Benefit the Poor?" *Economic Progress and Social Welfare*, Leonard H. Goodman, ed. (New York, 1966), pp. 125-57. Also available as Reprint No. 6, Institute for Research on Poverty.
10. ———, "Transfer and Redistribution as Social Process," *Social Security in International Perspective*, Shirley Jenkins, ed. (New York, 1969).
11. ———, "Toward an Economics of Health, Education, and Welfare," *J. of Human Resources*, Summer, 1966, pp. 43-53.
12. IRENE LURIE, "Transfer Payments and Income Maintenance," mimeographed Staff Paper for the President's Commission on Income Maintenance Programs (1969).
13. NELSON D. McCLUNG, "Problems in the Development of Data Bases for the Static Microsimulation of Income Transfer Program Direct Effects," mimeographed Staff Paper for the President's Commission on Income Maintenance Programs (1969).
14. IDA C. MERRIAM and ALFRED M. SKOLNIK, Social Welfare Expenditures under Public Programs in the United States (1929-66), Research Report No. 25, Office of Research and Statistics, Social Security Administration (Washington, 1968).
15. JAMES N. MORGAN, MARTIN H. DAVID, WILBUR J. COHEN, and HARVEY E. BRAZIER, *Income and Welfare in the United States* (New York, 1962).
16. MOLLIE ORSHANSKY, "The Shape of Poverty in 1966," *Social Security Bul.*, Mar., 1968, pp. 3-32.
17. The President's Commission on Income Maintenance Programs, Report of, "Poverty Amid Plenty: The American Paradox," preliminary copy, mimeographed (Nov. 12, 1969).
18. DOROTHY S. PROJECTOR, "Children's Allowance and Welfare Reform Proposals: Costs and Redistributive Effects," mimeographed paper prepared for meeting of the National Tax Association (Oct. 2, 1969).
19. ALVIN L. SCHORR, "Alternatives in Income Maintenance," *Social Work*, July, 1966, pp. 22-29.
20. JAMES C. VADAKIN, *Children, Poverty, and Family Allowances* (New York, 1968).

DISCUSSION

MICHAEL K. TAUSSIG: Professor Lampman's excellent paper first presents an accounting framework for analyzing the processes of extramarket redistribution in the United States and then goes on to discuss some current income maintenance reform proposals. I shall take up each of these quite distinct parts of the paper in turn.

The accounting framework Lampman outlines represents an attempt to classify and measure the processes of redistribution of incomes in much the same way the national income accounts treat the processes of production and the market generation of incomes. In any pioneering effort of this sort, difficult conceptual issues arise which require hard decisions and inevitably invite some disagreement as well. Among the several such issues that could be raised with Lampman's approach, I have selected two that especially interest me for discussion.

First, Lampman explicitly includes private insurance benefits related to health and income maintenance as part of the total extramarket redistributive process. He acknowledges the case for treating private insurance as purely market transactions but ends by including them in his redistributive framework on the grounds that (1) the insurance intermediary performs a unique service "in pooling risk and thereby converting factor income into payments that respond to stated contingencies rather than to accumulated contributions" and (2) "insurance benefits often flow to persons quite remote from the purchasers." I find these arguments most unconvincing. Consider the case of two individuals, A and B, exactly alike except that A uses a given portion of his net worth to purchase straight life insurance while B uses the same amount to purchase an option on a stock or on a site of land or whatever. Now if A and B die in the same accident, A's estate will include the face value of his insurance policy while B's estate will include the market value of his option contract. Why should the event of these two assets passing into the respective estates be treated differently in any redistributive accounting scheme? The sums of money payable to beneficiaries respond in both cases to stated contingencies rather than to accumulated contributions (and so do all gambling receipts). While A's insurance broker may perform a unique service, so may the real estate or stockbroker who makes a market in options (and so does the neighborhood bookie). Moreover, insurance benefits are not the only assets in an estate that may flow to persons quite remote from the purchasers. In short, I see no reason why private insurance benefits should be included in the Lampman redistribution accounting framework.

The second conceptual issue I wish to raise is of much greater substance and general interest. Lampman follows Gillespie in treating public education services as income-in-kind transfers to families with children in school. I doubt first of all whether the proper distinction between government services included or excluded in the accounting framework should have anything to do with some moot exclusion principle.

Pure, indivisible public goods can in some circumstances be important vehicles of redistribution. But the point I really wish to raise here is the implicit valuation of all government in-kind benefits by Gillespie and Lampman. In common with national income accounting practices, public services provided without explicit user charges are valued implicitly by Lampman at factor cost. However defensible this practice may be in measuring total product and income, it seems to me to be fundamentally wrong for the purpose of accounting for the redistribution of incomes. A strong case can be made, I believe, for the proposition that a dollar spent on the public education of a middle-class child is valued relatively high for both consumption and investment in human capital purposes by his family while a dollar spent on the public education of a poor child is valued relatively low by his family. Education for the middle-class child tends to be both a relatively pleasant and productive use of his time, while the poor child too often finds school both a painful and futile experience. In any event, to treat a dollar of government expenditure on education, health, or welfare services on a par with a dollar of outright cash transfers seems to me to be avoiding the very tough and fundamental issues involved in evaluating these nontransfer programs.

Finally, let me observe that Lampman's item 1, Table 1, "Preristribution incomes," should not be confused with the distribution of incomes that would result naturally in the absence of government intervention in the economy. On the contrary, all kinds of government activities directly and indirectly influence the composition of factor incomes in item 1, including the level and composition of government output and the resulting pattern of derived demand by the government for various factors of production. For example, one important class of government activity that importantly affects the distribution of factor incomes is regulation and licensing. To separate all the government activities that affect the distribution of incomes in item 1 from the activities that affect the "Postredistribution" pattern of incomes in item 20 is essentially arbitrary and defensible only on the grounds of convenience.

Let me comment now briefly on Lampman's generally excellent analysis of some income maintenance reform proposals. To my great dismay, Lampman apparently accepts the case for some kind of children's allowance program—at least under present circumstances. I object to the children's allowance approach under any circumstances because I cannot accept the underlying premise that the poor can be or should be separated into two distinct groups: deserving versus undeserving, employable versus unemployable, or some similar dichotomy. Such distinctions are nonoperational, if not immoral, so that the government is forced inevitably to make distinctions based on criteria like the absence of an able-bodied, working-age male in the family unit (AFDC) or the presence of one or more children in the family (FAP). The unfortunate poor who do not meet these criteria for eli-

gibility are shunted aside by some social pundits on the grounds of political feasibility; these undeserving or employable or just unlucky poor are urged to be patient and wait until the public becomes more enlightened about their plight. Such counsel reminds me very much of the similar advice given to black Americans ever since the Civil War; that is, patience and passivity until whites change their attitudes.

Beyond the broad, moral issues underlying the children's allowance controversy is the narrow technical question of the work incentive effects of a children's allowance program relative to the work incentive effects of a more comprehensive negative income tax. I hope that I read Lampman incorrectly when he apparently cites Alvin Schorr approvingly on this point. It simply is not true that a children's allowance scheme somehow avoids the static disincentive impact of a negative income tax because it is not tied directly to income. Any benefits payable under a children's allowance program have exactly the same income effect in discouraging work effort as do negative income tax benefits. Further, a nontaxable children's allowance benefit avoids negative income tax reduction of the benefits of poor families only at the cost of an enormous increase in marginal tax rates of families above the break-even point. The work-disincentive consequences of such a scheme could turn out to be most significant. As the rest of Lampman's discussion brings out very well, no income maintenance scheme can avoid the inherent trade-offs between adequate benefit levels with wide coverage of the poor, reasonably low marginal tax rates on the other income of the poor, and an acceptable net cost to be borne by taxpayers above the break-even point.

In conclusion, let me stress a point that Lampman touches on only briefly. A tremendous amount of redistribution takes place in this country within the family. In contrast, redistribution outside the family is manifestly inadequate. The generosity in our society within the immediate family coexists with incredible meanness outside the family unit. (Why this should be so is ultimately a question beyond what is now acceptable as economic "science," but it is interesting nonetheless to speculate whether this situation is necessary or inherent in a society based on private property and the efficient functioning of private markets.) But standing between the two poles of the nuclear family and all the "others" in society is the phenomenon of the "doubled-up" family. Some research indicates that most doubling-up is involuntary; that is, the sharing of a dwelling unit is a "bad" that individuals trade off against the economies of scale that can be achieved by such living arrangements. Most nonnuclear family combinations apparently will undouble if the income of each unit rises above very low levels. As Lampman indicates, the proper measurement of the family unit is basic to the study of income distribution; it is also basic to the design of income redistribution programs. Much work remains to be done in clarifying the significance of the nonnuclear family in this context.

T. PAUL SCHULTZ: Because of the fundamental so-

cial and economic importance of the personal distribution of income, I find it worrisome that economics has contributed so meagerly to our understanding of the causes and consequences of this outcome of our economic system. In reviewing two papers at this session I think some of the limitations of economic analysis in this field are illuminated and at the same time some promising directions for research and data collection are indicated.

To begin with Professor Thurow's paper, we have a regression analysis of the association between two parameters that summarize the size distribution of incomes and various aggregate indices of income, employment, inflation, etc. There is a long tradition for interpreting income distributions in terms of summary statistics that may be associated with inequality or poverty. Despite years of study by econometricians, statisticians, and other social scientists, no consensus has emerged on measurement in this field. The reason for this inconclusive discourse is, I think, quite clear: there is no value-free or theoretically appropriate conception of income inequality or poverty to measure. The two parameters beta distribution proposed by Thurow has two obvious shortcomings for this purpose: (1) neither parameter is uniquely related to inequality; and (2) the upper tail of the income distribution must be arbitrarily lopped off (at \$15,000) because the distribution has a fixed upper limit. Given these problems, Professor Thurow should marshal his evidence, in my judgment, to show that the beta distribution at least fits these truncated income data better than any of the many other functions already proposed in the literature by Pareto, Gini, Bortkiewicz, Gibrat, Simon, and Mandelbrot. Additional questions are also left unanswered in Thurow's presentation of his interesting findings. I am not certain, for example, whether the R^2 s presented in Tables 1 and 2 are derived from fitting the cumulated (Lorenz) or frequency distributions of income. What sort of nonlinear procedure is used to estimate the parameters of the beta distribution; is the search procedure sequential or simultaneous in the parameters? When changes in income "dispersion" and "concentration" are reported, I am curious how Thurow defines these terms, for they can apply to many statistics.

A second traditional issue in "economic" analyses of income distribution is the nature of the income recipient unit. Thurow does not address this issue, but deals with families and unrelated individuals together, distinguished only by color. Not only has the composition of income units, thus grouped, secularly and cyclically changed with implications for welfare (i.e., the doubling and undoubling of household units), but composition also differs substantially between white and nonwhite income units. A theory of family formation and labor force participation would appear to be an essential tool if we are to infer anything precise about changes in welfare from changes in income, distributed among a changing composition of income units.

Despite these reservations, the regression analysis finds associations of a plausible sort. The strength of the relationships undoubtedly turns on the choice of

distributional parameters, for the parameters reflect the secular growth of income levels during the period, which is extenuated by the neglect of incomes over \$15,000 (which constitute at least 10 percent of total personal income by the end of the period). Thus, it is hardly surprising that where "pure" or relative measures of income inequality, skewness or dispersion exhibit erratic associations with secular trends in productivity and price level, stronger associations are found when the predicted parameters of the income distribution contain substantial information on both the level and dispersion of incomes.

Further aspects of Thurow's model deserve comment. The two parameters of the beta distribution that jointly determine the degree of income "dispersion" or "concentration" are very highly correlated: in the time series the simple correlation between ρ and σ is .97 for whites and .84 for nonwhites. It is therefore plausible to expect that the disturbances in one equation will not be independent of those in the other. Evidence of such correlation between the disturbances suggests that the two equations are part of a simultaneous system and at least more efficient methods of estimation than ordinary least squares could be used.¹ Moreover, given prior knowledge of the relative stability of pure measures of income inequality during the postwar period, an argument could be advanced for estimating the two parameters of the beta distribution as part of a recursive model.² Indeed, the computer procedure used for estimating the parameters of the beta distribution may itself be recursive or sequential in its optimizing algorithm.

There is finally the problem of collinearity between explanatory variables or the nearness to singularity of their correlation matrix.³ In the presence of severe collinearity, as is probably the case in Thurow's time series, parameter estimates of the model are likely to be sensitive to which variables are excluded from the regression. Thurow's practice of dispensing with all explanatory variables for which the t statistic is less than one is open to challenge, given the interpretation he proceeds to make of the relative magnitudes of his parameter estimates. First, it implies one has no a priori specification of the model in mind, calling into question the meaning of classical hypothesis testing. Second, unless one demonstrates that the relative magnitudes of the final parameter estimates are insensitive to the inclusion of "less significant" explanatory

variables (those excluded in the reported results), the parameter estimates hinge precariously on the final model specification. It is asking rather much of data to require them to tell us of the causal structure by which they were generated.

In summary, Thurow's regression findings are statistically significant by conventional standards. But are these standards appropriate when no explicit analytical framework is set forth, and no clear hypotheses dictate the specification of the model, the choice of the data, or the appropriateness of the estimation procedure? I judge not.

Professor Budd has sought to refine and improve the descriptive data base on which we rely for evidence of postwar change in the size distribution of personal income. He is keenly aware of the deficiencies of these data, the differences among data sources, and their changes over time. The apparent stability of aggregate income inequality, measured by such a pure summary statistic as the Lorenz coefficient, has led Budd to design a computer routine to smooth and refine these various data sources to discover what perceptible underlying shifts in income shares and relative mean income levels have occurred since the second World War. His new evidence from these relatively independent data sources (e.g., CPS, IRS, CES) confirm that beneath the overall stability of income inequality, the lower 40 percent of the income units and the upper tail (top 4 to 5 percent) have lost ground on the relative income scale, whereas the top 50 to 95 percent of the income units have improved their relative income status. This confirms the less refined evidence he initially presents from published OBE estimates.

What can we conclude from Budd's new evidence of small but systematic changes in the size distribution of personal income? I am doubtful whether one will be able to draw strong conclusions about the distribution of economic resources among persons until one has a theory of how the composition of income units changes under cyclical economic and demographic pressures. Though Budd's findings are still fresh from the computer, and as he indicates, as yet incompletely digested, I would urge him to pursue them beyond the descriptive plane and search for testable economic hypotheses concerning the determinants of changes in relative shares of high-, middle-, and low-income units.

This brings me back to my initial inquiry: why has economics shed so little light on so important a question as the causes and consequences of the personal distribution of income, inequality or poverty? The lack of a macrotheory is central. At the micro level one has a relatively firm set of notions concerning the determination of personal income. From much simplified assumptions concerning the interaction of a few attributes and income and their correlations, some deductions have been offered about income variances and skewness.⁴ But the growing wealth of mi-

¹ Undoubtedly greater efficiency could be gained by the estimation procedure proposed by A. Zellner, "An Efficient Method of Estimating Seemingly Unrelated Regressions and Test for Aggregation Bias," *J. Amer. Statist. Assoc.*, 1962, pp. 348-68.

² If a recursive or jointly determined simultaneous model were assumed, in which say σ were estimated as a function of a given or endogenously determined value of ρ , different parameter estimates for the remaining explanatory variables than Thurow's would be anticipated.

³ See, for example, Yoel Haitovsky, "Multicollinearity in Regression Analysis: Comment," *Rev. of Econ. and Statis.*, 1969, pp. 486-89.

⁴ Gary S. Becker and Barry R. Chiswick, "Education and the Distribution of Earnings," *A.E.R.*, May, 1966, pp. 358-69.

crodata has not been fully exploited to construct the analytical bridge for aggregate analyses of income distribution based on the empirically tested assumptions embodied in microeconomic models of personal income determination.⁶ I hope the distribution of economic professional talents allocated to this task shall increase in the future.

STEPHAN MICHELSON: "Political economy begins with the fact of private property without in any way accounting for that fact. Instead, it takes the material processes of private property as they actually occur, presents these processes in general, abstract formulas, and then offers these formulas as laws." This is surely not the first time Karl Marx has been quoted at a meeting of the A.E.A., but it must be one of the most appropriate. Lester Thurow's paper offers abstract formulas which in no way account for the mechanism by which the phenomenon under study, income distribution, occurs. This basic criticism, in my view, would vitiate any otherwise valid findings. Since the methods employed are insufficient even for the limited purposes of this paper, I will start with a short, technical critique. The basic problem will not be, however, that this study is weak econometrically, but that it is weak conceptually. Nor, harsh as my remarks may appear, are they really directed at Thurow, or his efforts. They are directed at the economics profession which encourages and praises such misdirection of effort and, in this case, of obvious ability.

Time Series Analysis. From his equation (3), Thurow will try to determine the effect of government purchase of goods and services (the exponent of GP/H), while holding GNP/E constant. If he has a model of the economy in mind, it apparently includes a multiplier of 0. Full employment—with only a monetary increase in GP/H —is not assumed, because I , the inflation index, is also held constant. Later reference to the Phillips curve cannot change the system implied by the regression estimation. Similarly, when GNP/E is conceptually incremented, PI/GNP is held constant, which means PI must increase in proportion to GNP , a restrictive condition in an income distribution study, and not true over the business cycle as Thurow himself notes. In short, the use of a single regression on these variables cannot answer the questions asked. The income distribution effect of an increase in government purchases should take account of the change in PI/GNP induced, as well as the change in prices and/or

⁶ One promising method of extracting information from large files of individual data about the incidence, persistence, and control of poverty is to treat the phenomenon as probabilistic and use the theory of Markov processes to describe movements of the population into and out of poverty. This methodology has been proposed and applied to social security longitudinal data by J. J. McCall in two studies: *An Analysis of Poverty: A Suggested Methodology*, RM-5739-OEO (Oct., 1968), and *An Analysis of Poverty: Some Preliminary Findings*, RM-6133-OEO (Dec., 1969, RAND Corp., Santa Monica, Calif.).

⁷ Thurow interprets GNP/E as a "growth" variable, when clearly this measures the level, and $d(GNP/E)/dt/GNP/E$, the proportionate time rate of change of GNP , would measure growth.

real income. If PI/GNP is a function of GNP/E , or I is a function of E/LF , etc., they should be estimated that way. Total differentiation is called for, which is to say some model of the economic system, including its interactions, and an appropriate estimator of that system.

Given that the estimates were not derived from an underlying model, interpreting them is a problem. Consider the following example: The equation for the white ρ parameter from Thurow's Table 1, written in the form of his equation (3), looks like this:

$$(1) \quad \rho = 3.97 \left(\frac{GNP}{E} \right)^{1.30} (I)^{-1.55} \left(\frac{PI}{GNP} \right)^{-2.43} \cdot \left(\frac{E}{LF} \right)^{-2.52} \left(\frac{TP}{H} \right)^{.49}$$

Assigning the exponent to both the numerator and denominator of each expression, and combining similar terms by adding the exponents, gives the following:

$$(2) \quad \rho = 3.97 \left[\frac{(GNP)^{3.63}(LF)^{2.52}(TP)^{.49}}{(E)^{3.72}(PI)^{2.43}(I)^{1.55}(H)^{.49}} \right]$$

From the standard errors in Table 3, it is obvious that the exponents of GNP (3.63) and E (3.72) are not significantly different; nor are the exponents of LF (-2.52) and PI (-2.43). Equation (1) (above) is therefore equivalent (using midpoints of these insignificant differences) to

$$(3) \quad \rho = 3.97 \left(\frac{GNP}{E} \right)^{2.67} (I)^{-1.55} \left(\frac{PI}{LF} \right)^{-2.47} \left(\frac{TP}{H} \right)^{.49}$$

This equation has one less term than Thurow's, to some extent reducing both redundancy in the variables and the resultant difficulty of interpretation. Of course I am not claiming that the exponents in my equation (3) would have resulted from direct least squares estimation; only that this is an algebraically equivalent form of Thurow's estimates. The interpretation, however, is now completely different. His conclusion about the effect of "growth" reads: "Since the income dispersing effect of ρ is roughly twice as large as the income concentrating effect of σ , the income dispersing effects cancel each other. . . . Growth . . . does not have any major impact on the dispersion of income." (Emphasis mine.) Multiplying my (GNP/E) exponent by two, and comparing it to the (GNP/E) exponent for σ , the italicized part should read: ". . . the income dispersing effect is to the concentrating effect as 7.34 is to 2.94, or essentially 2½ times as strong. Growth tends to produce inequality in white incomes."

Cross-section Analysis. The cross-section equation, curiously unrelated to the time series equation, contains analogous errors. Starting with interpretation: Thurow's Table 2 does not provide substantiation for his statement that "being on a farm or in a family with no one in the labor force is not the relative handicap [for blacks that] it is for whites." The same state-wide data⁸ were used as observations for blacks and whites.

⁸ I am not commenting in this note about the data employed in either the time series or the cross-section regressions. However, weighting states by their population, N , whatever the data source, is not necessarily

Therefore one might say that being in a state with relatively more families with no one in the labor force, or relatively more families on farms, does not affect the income distribution of blacks within that state. Thurow's statement is about families; mine, about states. Both statements may be true, but the equation relates only to the latter. This being the only kind of statement which can result from these state-wide observations, it is not clear to me what the point of the regression is.

Since the kind of work offered to blacks is low paying and less than full time, it is something of a surprise to find Thurow advocating high-wage industries and full-year jobs. In fact, wages and full-time employment are correlated. Their joint effect cannot be estimated without an interaction term.³ Even if it were technically correct, this kind of conclusion is similar to that from correlating percent of blacks living in an area with those working in that area, finding that blacks tend to live near where they work, and concluding that if blacks were residentially relocated, they would earn more.⁴ The important question in both cases would seem to be whether the blacks would in fact be hired.

On Studying Income Distribution. Thurow justifies his equations because they work, but feels free to interpret them as if they constituted a fully specified economic model. What is missing throughout this study—and in virtually all simplistic econometric studies of this issue—is an attempt to ask what is happening, what the income generating process is.⁵ Employment and income are the results of a whole social system, not likely to be described by a single equation, and not likely to be estimated by ordinary least squares. This system might be so structured as to tend to produce a kind of income distribution which is characterized by inequality. Further—

correct. Firms are not usually weighted in production function estimation, though to correctly portray the average effects of inputs they should be weighted by output. If each state is thought to produce a ρ or a σ , it might be considered an independent (equal-weighted) observation. If, on the other hand, one wanted to correct for error in the sampling, which is presumably proportional in each state to that state's population, then weighting by \sqrt{N} would be appropriate.

³ Where Y is earnings, W is wages, and H is hours,

$$\Delta Y = W\Delta H + H\Delta W + \Delta W\Delta H.$$

The neglect of such interaction can lead to serious error. In commenting on Thurow's absurd finding that blacks' lack of education produced over 40 percent of the black-white income gap, I pointed out that interaction between race and education was at fault. See Lester C. Thurow, "The Causes of Poverty," *Q.J.E.*, Feb., 1967, pp. 39–57, and Stephan Michelson, "Income of Racial Minorities," Chap. 4 (Ph.D. dissertation, Stanford Univ., 1968).

⁴ It does seem ridiculous, doesn't it? See John F. Kain, "Housing Segregation, Negro Employment, and Metropolitan Decentralization," *Q.J.E.*, May, 1968, pp. 175–97.

⁵ For an exception to this rule, see David P. Taylor, "Discrimination and Occupational Wage Differences in the Market for Unskilled Labor," *Ind. and Labor Rel. Rev.*, Apr., 1968, pp. 375–90.

more, and more importantly, the system might regulate the placement of people along the income distribution.

I do not mean that one's future place on the income distribution or the shape of that distribution is fixed at birth. But there are certain well-known rules to making income in this country: be white, have "middle-class" values (punctuality, diligence, courtesy, etc.), know whom to flatter, whom to scorn.⁶ In general these qualities are correlated with type of family at birth. But this relationship is not inexorable. Though anyone who can play by the rules can win (Richard Nixon, for example), it is important to see that these rules are not neutral: they favor certain classes. The most important of these is the propertied class. The most ignored is the property-less nonwhite (though the most actively persecuted is the American Indian).⁷ Certainly Lester Thurow does not think that business will give high wages and full employment to blacks because he says that is how blacks can raise their incomes. What was the point of expending his valuable intellectual resources on such an unproductive task?

My major objection to this type of study, then, is that it cannot, by design, ask serious questions. Such questions would have to probe the nature of the system which produced the results which are obvious in the data. Even the correlation between ρ and σ over time for whites, but their lack of correlation for blacks is not explained. I get absolutely no feel for the social and economic decision making, and the institutions which direct that decision making, the results of which are the beta distribution parameters. Nor are the institutions which allocate high and low rank, whatever the distribution, considered. Economists' "policy pronouncements" tend to be as hollow as their analyses: from income regressions on industrial structure and weeks worked, what could one say except "improve the industrial structure," "work more weeks"? The analysis limits the conclusions to the trivial.⁸

⁶ There is evidence on the extent to which these rules are institutionalized. Bernard Mackler ("Grouping in the Ghetto," *Educa. and Urban Soc.*, Nov., 1969, pp. 80–96) reports about a Harlem school "that the price for success is behaving in a 'school' (or socially) acceptable way. The school staff rewards pupils who are passive, who listen, who are polite, and who adhere to the rules." For a compilation of evidence on the characteristics of workers which are rewarded by the economy, see Herbert Gintis, "Alienation and Power: Towards a Radical Welfare Economics" (Ph.D. dissertation, Harvard Univ., 1969).

⁷ Recent econometric work by Michael Reich at Harvard suggests that the distribution of income among whites is related to the extent of racial discrimination against blacks. Where, in Thurow's equations, does the political-economic issue of discrimination enter?

⁸ Since a government policy to actually equalize black and white incomes would not be acceptable in this system, I stressed analysis which might lead to black policy in, "On Income Differentials by Race: An Analysis and a Suggestion" (*Conference Papers of the Union for Radical Political Economics*, Dec., 1968). The socially concerned economist might consider this course: address questions to the policies and actions of those people in whose interest they would work.

Thus, to repeat, my criticism is not to be interpreted as focused on this paper alone. In a flu epidemic, one sneeze more or less is not too important. Nor do I imagine that one vial of antibiotic, even placed at the center of the epidemic, will seriously affect the spread of the germs. But perhaps we can start talking about

the kinds of questions we ask in economics, about the extent to which we have failed to investigate the impact of the basic social, political, and economic system in studying phenomena such as income distribution. Then Marx will have had his day, and we will all be healthier for it.

DYNAMICS OF INCOME DISTRIBUTION: POVERTY AND PROGRESS

VARIABILITY OF ECONOMIC WELL-BEING AND ITS DETERMINANTS*

By JAMES D. SMITH, *Pennsylvania State University*
and

JAMES N. MORGAN, *University of Michigan*

Introduction

What changes occur in family economic well-being over time and what are some of the determinants of these changes? To shed light on these questions, an ongoing panel study of some 4,500 families will shortly begin its third wave of interviews.¹ Data gathered from the first two waves provide the variables analyzed in this paper.

Section I presents evidence that for a large proportion of families there is a substantial change in money income from one year to the next. The same holds true for a measure of well-being which is expressed as a ratio of family money income to family needs. When this ratio is equal to one, it approximates the Orshansky poverty line.² The second section analyzes the importance of some of the components of change in money income and well-being. Finally, in the third section we look at the more fundamental attitudinal and behavioral characteristics which determine change in well-being.

I. Changes in Income and Well-Being

As we measured it, money income includes wages and salaries, mixed wage-capital income, asset income, and regular money transfers. It does not include capital gains, realized or unrealized.

Table 1 presents a matrix showing the per-

* The valuable assistance of Jonathan Dickinson is gratefully acknowledged. He offered valuable criticism and did yeoman labor in the computer running of the data.

¹ A general description of the design of this study was presented a year ago at these meetings. See James N. Morgan and James D. Smith, "Alternative Measures of Well-Offness and Their Correlates," *A.E.R.*, May, 1969. Also see *A Longitudinal Study of Family Economics*, Inst. for Soc. Res., Univ. of Michigan (free on request from the authors). For a more detailed treatment of the study see James N. Morgan and James D. Smith, *A Panel Study of Income Dynamics* (Inst. for Soc. Res., Univ. of Michigan, 1969).

² See Mollie Orshansky, "The Shape of Poverty, in 1966," "The Poor in City and Suburb," and "Who's Who Among the Poor, A Demographic View of Poverty," appearing, respectively, in the *Soc. Sec. Bul.* of Mar., 1968, Dec., 1966, and July, 1965.

tage distribution of families by 1967 and 1968 money income decile for all families who were interviewed in both 1968 and 1969. Reading down the diagonal it is apparent that there is a considerable amount of churning about within the income distribution. Looking at the 1967 interdecile range, 60 percent to 70 percent of the family units in each decile rank were no longer in that rank in the distribution of 1968 family income.

Even at the two extremes of the diagonal, a large percentage of family units moved out in the one-year interval. About 30 percent of the families in the lowest decile of 1967 income had moved to higher deciles in the rank of 1968 income, and of those moving, over 40 percent shifted two or more deciles. The rich, too, lose their grip on income rank. About 36 percent of these family units dropped out of the most affluent 1967 income decile. About one-half of those who dropped slid down one decile, but the other half of the descenders are strung out across the lower deciles of 1968 income.

Two factors may explain these downward shifts of high-income families. They receive proportionally more income from self-employment than do low-income families, so that the vicissitudes of business and farm operations may account for larger swings in family income. Second, the way we treated split-offs in the matrix—mainly sons and daughters who left home and husbands and wives who separated between the two interviews—may have resulted in some large movements. When an adult member left his 1968 family, both the original family and the departed member were sought for interview. In the processing of these interviews, 1967 income of the original family unit was related to the 1968 income of each family unit resulting from the split-up. Especially in the case of children leaving home, the difference in income could be quite large.

Despite the substantial amount of movement out of both the top and bottom 1967 income deciles, about two-thirds of the units in those deciles did in fact remain there in the rank of 1968 income, while only about one-third of the families

TABLE 1
DECILE RANK OF FAMILIES BY TOTAL MONEY INCOME, 1967×1968

Decile Rank 1967 Income	Decile Rank 1968 Income										%
	\$2,199 0	\$3,619 1	\$5,069 2	\$6,499 3	\$7,874 4	\$9,399 5	\$10,999 6	\$13,099 7	\$16,859 8		
0	69.2%	17.7%	4.5%	4.3%	2.7%	.4%	.6%	.4%			100.0
\$1,976											
1	19.2	43.9	20.6	8.4	3.5	1.6	1.4	.7	.7%	.1%	100.0
\$3,219											
2	6.7	22.5	37.2	15.3	8.4	5.2	2.0	2.1	.1	.5	100.0
\$4,649											
3	1.9	5.6	17.8	38.1	21.2	7.4	4.8	1.1	.6	1.5	100.0
\$5,999											
4	1.5	5.7	9.6	19.6	29.7	19.3	7.6	3.9	2.5	.6	100.0
\$7,206											
5	1.4	2.3	3.6	8.1	19.2	34.6	22.4	5.1	1.8	1.6	100.0
\$8,602											
6	.7	1.3	2.7	3.2	7.8	19.1	34.9	20.3	7.7	2.3	100.0
\$10,175											
7	.7	.7	.3	.7	4.2	4.9	16.9	40.9	22.6	8.1	100.0
\$11,999											
8	.7	1.0	2.3	2.0	2.4	3.6	8.6	19.9	43.3	16.2	100.0
\$15,389											
9	.6	1.2	1.9	1.7	2.5	5.1	1.4	4.6	17.5	63.5	100.0

NOTE: Only families from whom we were able to obtain an interview in the spring of 1968 and 1969 were used in this matrix. The decile rank for 1967, however, was computed using all families who gave an interview in 1968, even if they did not give a 1969 interview.

in the interdecile range of 1967 incomes stayed put. Part of the greater stability of the extreme deciles is attributable to their open-endedness. But in addition, a disproportionate number of families at the very bottom of the income distribution are headed by a female with children or an aged or disabled person—all people for whom income-increasing opportunities are limited. On the other hand, high levels of education and the stability of interest and dividends in the incomes of the affluent probably operate to retain high-income families in the top decile.

The same pattern of change but in a more exaggerated form exists when we move to the more complex, Orshansky-type measure of well-being. This measure is similar to that used by the Social Security Administration and the Office of Economic Opportunity and takes account of both family income and family needs. We calculated the minimal food cost for each family member according to his age and sex and summed them to arrive at a total family food cost. The total family food cost was then adjusted for economies of scale and the result multiplied, in turn, by a factor based on family size, to arrive at a total need standard. The need standard was then divided

into the family's reported income to arrive at the Orshansky score.³

Since our Orshansky measure of well-being is a ratio, a family's Orshansky score may change because of a shift in either money income or family need. Changes in need come about from alterations in family size, aging of family members, and shifts in the head's occupation from farm to non-farm categories.⁴

The matrix of families by Orshansky score in Table 2 was constructed in the same manner as the matrix by income described above. The 1967 Orshansky score is based on 1967 family income and the 1968 score on 1968 income.

Although the pattern of movement of families across deciles is similar to that shown for money income, it is apparent that families change their relative position in the distribution of Orshansky

³ See appendix for the food cost scale and adjustment factors used. Space allocation in this issue of the *A.E.R.* did not permit publication of the appendix. It is available from the authors.

⁴ Families whose head is employed on a farm had their needs standard reduced by 20 percent on the grounds that the money cost of living for farmers is lower than for nonfarmers.

TABLE 2
DECILE RANK OF FAMILIES BY ORSHANSKY RATIO 1967×1968

Decile Rank 1967 Orshansky Score	Decile Rank 1968 Orshansky Score										%
	.81 0	1.20 1	1.56 2	1.94 3	2.32 4	2.71 5	3.23 6	3.88 7	4.89 8	9	
.74	0	61.9%	19.5%	8.3%	4.4%	2.2%	1.9%	1.0%	.7%	.1%	100.0
1.09	1	23.1	40.8	18.2	8.6	3.6	3.2	.7	.9	.6	.3% 100.0
1.45	2	7.1	22.0	32.1	18.5	8.2	5.5	2.8	1.3	1.4	1.1 100.0
1.78	3	2.0	7.8	20.1	30.6	21.7	7.1	6.4	2.9	1.0	.3 100.0
2.13	4	1.6	4.0	9.2	18.6	29.6	21.1	10.0	3.1	1.5	1.3 100.0
2.54	5	1.0	1.7	4.7	8.7	19.9	31.6	20.4	8.8	2.1	1.1 100.0
3.04	6	.7	1.5	2.8	5.2	6.6	16.1	32.7	24.1	8.1	2.1 100.0
3.66	7	1.3	1.5	1.5	1.9	4.0	8.5	19.7	33.6	22.1	6.0 100.0
4.60	8	.3		1.6	1.1	3.1	2.8	4.9	18.0	47.1	21.1 100.0
	9		1.4	1.2	1.8	1.0	2.4	1.4	7.0	16.8	66.8 100.0

NOTE: Only family units from whom we were able to obtain an interview in both the spring of 1968 and the spring of 1969 were used in this matrix. The decile rank for 1967, however, was computed using all families who gave an interview in 1968, even if they did not give a 1969 interview.

score more frequently than they do their relative income position.

II. Components of Change

Taking total family income as the sum of the labor income of the head, wife, and other family members, plus income from capital and transfers, it is of interest to know how changes in each of these sources affect changes in total family income. For this analysis we selected only those families for whom we had observations in both interviews and who had the same head in both years.

The intercorrelations among the changes in

components and the change in total family income are shown in Table 3.

As would be expected, changes in the head's income are correlated more closely to changes in total family income than are the other sources, but even changes in head's income explain less than 45 percent of the variance in the change in family income. Changes in labor income of the wife were also found to be moderately correlated with changes in family income, but no correlation appeared between changes in wife's labor income and head's labor income. A negative correlation was expected in the latter case on the grounds that working wives play a defensive role, moving

TABLE 3
CORRELATION COEFFICIENTS AMONG INCOME SOURCES

	1	2	3	4	5
1. Δ Total family income.....					
2. Δ Head's labor income.....	.67				
3. Δ Wife's labor income.....	.38	.01			
4. Δ Other family member labor income....	.38	.00	-.01		
5. Δ Family capital income.....	.27	-.14	-.02	-.02	
6. Δ Family transfer income.....	.14	-.22	-.02	-.07	-.02

into the labor force when the husband's income shifts downward and moving out when husbands receive large increases in income. It appears from the coefficients, however, that such behavior does not occur in a one-year period.

With respect to the Orshansky score, $O = Y/N$, we are interested in how changes in the family needs standard and income affect the ratio. Table 4 shows the intercorrelations among the ratio and its components. As one would expect, ΔO is strongly correlated, $r = .86$, with ΔY , but surprisingly weakly correlated with ΔN , $r = -.07$. The latter can be attributed in part to the weak positive correlation of ΔN and ΔY , $r = .15$. One possible explanation is that family size increases most during the same years when the income of the family head is moving up most rapidly. It also may be the case that the spur of greater family responsibilities pushes family heads to seek more income.

III. Determinants of Change

Short-run variations in individual family income occur because of abrupt structural change in the family itself; slower acting motivational forces, abilities of family members, and external economic shocks.⁶ Prime examples of family structural change are shifts from full-time employment to retirement of family members, or return of a wife to the labor force after raising children. Motivational factors are characterized by levels of aspiration, time horizon, and risk avoidance. Years of education may serve as a proxy for ability or level of human capital. External economic shocks include strikes or layoffs, change in local labor demand, and accidents or illnesses of bread winners.

One of our concerns in this study is the importance of motivational factors. To assess their impact we selected a sample of families in which we believed structural changes and external shocks had not been strong enough to mask changes in income resulting from attitudinal forces. Thus only families in which the same head was interviewed in both years were used. This eliminated

⁶ A fifth factor which gives the illusion of variation in income for individual families is measurement error. We have only two observations on each family and the combined measurement error may represent a sizable portion of the observed change in income for a single family. Even if, in fact, there were no change in income for any family but there was a normal distribution of measurement error around each observation, it might appear that a great deal of change was taking place. This would occur because only a small portion of the paired observations would fall either on the true level of income or on deviations from the true levels which were completely compensating.

TABLE 4
INTERCORRELATIONS AMONG SELECTED VARIABLES

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Aspirations.....	—													
2. Efficacy.....	.15	—												
3. Trust-hostility.....	.08	.23	—											
4. Risk avoidance.....	.01	.25	.21	—										
5. Disability.....	.33	.24	.11	.14	—									
6. Current handicaps.....	.21	.25	.15	.26	.58	—								
7. Background problems.....	.20	.13	.13	.20	.20	.24	—							
8. Education.....	.34	.28	.21	.36	.30	.39	.34	—						
9. Age.....	.57	.14	.07	.14	.36	.21	.12	.43	—					
10. Income 1967.....	.30	.24	.18	.33	.30	.31	.26	.14	.15	—				
11. Δ Income.....	.12	.09	.05	.03	.10	.08	.05	.39	.00	.06	—			
12. Orshansky 1967.....	.13	.21	.17	.37	.25	.29	.19	.13	.11	.83	.17	—		
13. Δ Orshansky.....	.07	.07	.04	.03	.06	.04	.01	.02	.07	.86	.16	.15	—	
14. Needs.....	.34	.03	.02	.13	.14	.03	.10	.02	.28	.31	.13	.04	.07	—
15. Δ Needs.....	.04	.00	.02	.04	.07	.01	.04	.01	.09	.09	.15	.07	.02	.02

most units in which death or change in marital status caused a change in structure and also tied the attitudinal measures to the family head. Second, families which received more than 10 percent of their income from transfers or were headed by a person over age 75 were excluded, thereby eliminating ADC families and retirees, whose incomes are constrained. Finally, families in which the head worked fewer than forty weeks in either 1967 or 1968 were eliminated, to avoid those large changes in income which occur because of major changes in employment status.⁶

The remaining sample of 1,274 family units had a mean change in income of \$984, with a standard deviation of \$2,264. Taking this change in income as the dependent variable, we utilized a computer algorithm known as AID (Automatic Interaction Detection Program)⁷ to identify those predictor variables which exerted the greatest influence on income change. Although our focus was the impact of certain attitudinal variables on income change, included in the list of independent variables were certain demographic and situation variables which might impose important facilitating or restricting conditions.

The independent variables employed are listed in Figure 1. The attitudinal measures are indices built up from reported behavior that took place over the year, family conditions at the time of the interview, and responses to direct attitudinal questions.

Figure 1 is a schematic diagram of the AID algorithm as it ran on our sample of "main stream" American family units. The first division comes on family size, separating one- and two-person families into one group with a mean change in income of \$493 and larger families into a second group with a mean change of \$1,193. The influence of family size on change in family income appears to be a structural factor; that is, large families are more likely to have additional

members entering the labor market in any period of time than smaller families. Thus the mean change in income of only \$984 for the sample as a whole could be exceeded severalfold by the increment to family income produced by a son or daughter who was a student in the first year but employed in all or part of the second.

For small families (left side of figure), age and education are the only variables which significantly increase our ability to predict change in income. For small families with heads 55-74 years old, no further division on any of our independent variables could increase the explained variation by .35 percent of the original sum of squares. In the case of small families with heads under 55, a substantial difference in mean income change occurs between those who had a college degree and those who did not, \$1,982 compared to \$550.

Families of three or more persons (right side of figure) with heads who scored high on our index of risk avoidance had a much higher average change in income than did those with heads who scored low, \$1,370 as opposed to \$881. And among risk avoiders, families of five or more persons had a larger mean increase in income than did three- and four-person families. Finally, high aspirations and a college education interacted to yield the largest mean increase in income, \$2,312, among large families headed by risk avoiders. The smaller families headed by risk avoiders fared substantially better if the head also had a high score on efficacy and planning, \$1,300 compared to \$444. And those with a high-efficacy score split on age, heads over 35 having a larger average increase in income at \$1,554 than those under 35 at \$977.

On the other hand, families of three or more persons with heads who were risk takers also did well if they were headed by a person who scored very high on our scale of efficacy and planning. They had a mean increase of \$1,618, which compares very favorably to the increases of other terminal groups. The risk takers with a low sense of efficacy enjoyed larger average increases in income, \$1,027, if they were headed by a high school graduate than if they were not, \$451.

In summarizing these results it is interesting to observe that education, where it is significant, places last after the middle-class virtues of aspiration, efficacy and planning ahead, and risk avoidance in enlarging the predicted value of change in family income over the year. It should also be noted that sex, race, disabilities, background problems, trust-hostility score and current handicaps could not contribute significantly to a reduction of variance.

Despite the consistency with which the attitu-

⁶ Also excluded were cases in which more than 10 percent of the family's income in either year had to be imputed during the editing of questionnaires and any additional cases in which the change in income exceeded three standard deviations.

⁷ See John A. Sonquist and James N. Morgan, *The Detection of Interaction Effects* (Inst. for Soc. Res., Univ. of Michigan, 1964). The algorithm begins by computing for the full sample the mean and sum of squares of the dependent variable (change in family income in this case). It then selects from among a set of independent variables the one which can split the sample into two groups with the smallest combined sum of squares. Each of the resulting groups is then treated as an initial sample and the process continues through series of binary splits. In our analysis we confined ourselves to splits which reduced the total variance by .35 percent and did not result in groups of fewer than twenty cases.

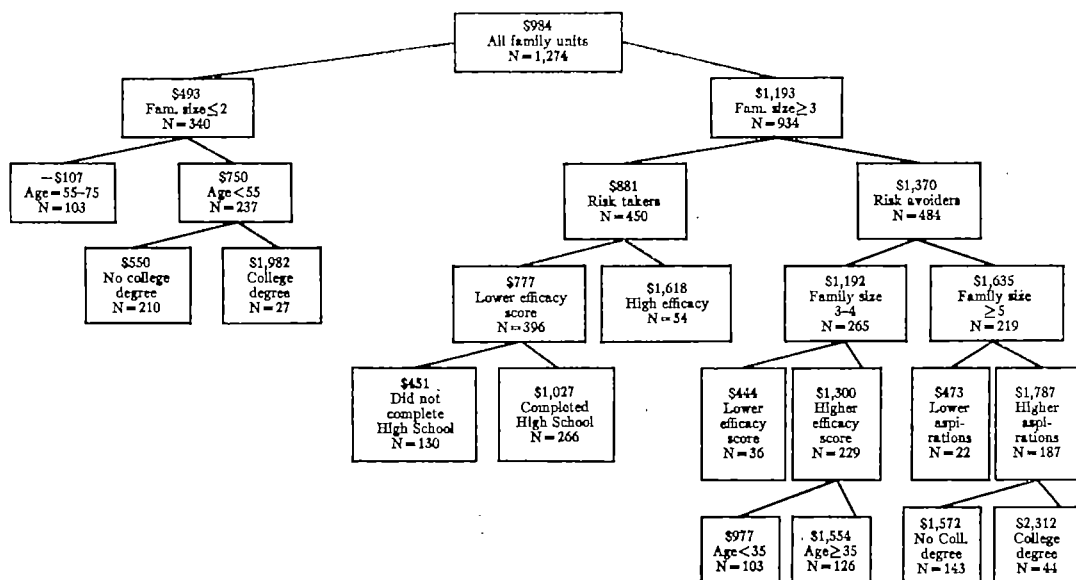


FIGURE 1

CHANGE IN MONEY INCOME FOR FAMILIES WITH WORKING HEADS*
(Top figure in each box is mean change for group)

* The sample of families used was restricted to those with the same head interviewed in both years, who received no more than 10 percent of their income from transfers, and in which the head worked full or part time 40 or more weeks in each year. Families were also excluded if more than 10 percent of their income was assigned, or if their reported change in income exceeded three standard deviations of the change for all families in the full sample of 4,460 units. The 11 terminal groups explain about 11.5 percent of the sum of squares in the total sample, allowing for two splinter groups which split with extreme values and N's of 3 and 7, which are not shown. Variables used in the analysis are listed below. Only those shown in the diagram above were able to reduce significantly the unexplained sum of squares in the original group.

Age of head	Family size	Unemployment rate in county	Efficacy and planning
Sex of head	Race	Current handicaps	Trust-hostility
Education of head	Disability of head	Expressed aspirations	Risk avoidance
			Time horizon

dinal variables related to income change, only about 11.5 percent of the total variance in income change was explained by the analysis. This suggests that external shocks and structural changes not accounted for by the family size variable may play an important role in determining short-run changes in family income. Nevertheless, because shocks tend to be random and structural changes may even out over time, observations over several years may show that attitudinal variables are of even greater importance than indicated by our present results.

Families at the bottom of the income distribution are of special interest because of our society's expressed commitment to ameliorate the hardships of the poor. We extracted from our sample all those families with the same head in both years, who were in the lower three deciles of the 1967 income distribution, and tried to explain

changes in their income between 1967 and 1968 with the following set of variables:⁸ family size, sex of head, race, employment status in 1968, age of head, background problems, current handicaps, aspirations, time horizons, county unemployment rate, efficacy and planning, trust-hostility, risk avoidance, change in family work (hours and number employed). The results of the AID algorithm to reduce the sum of squares around the mean change of \$935 for all families in the bottom three deciles are shown in Figure 2.

Of major importance to the size of the change in poor family income was the simple structural factor of 1968 employment of a wife who was not

⁸ Families were excluded if the head was over 75, they had more than 10 percent of their income assigned in editing the questionnaire, or if the change in their income was beyond three standard deviations of the mean change for all families.

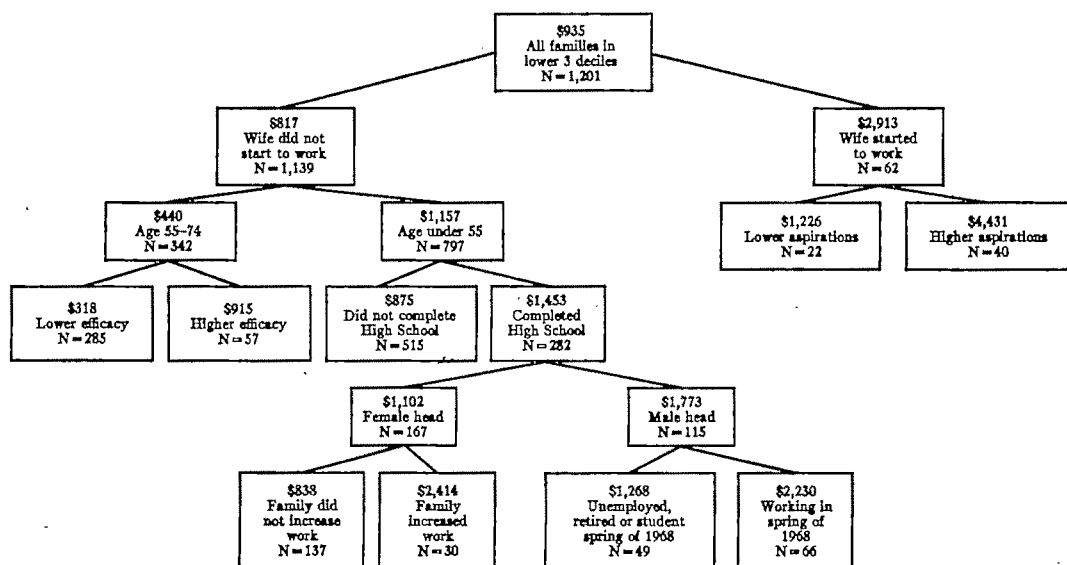


FIGURE 2

CHANGE IN INCOME OF FAMILIES* IN THE BOTTOM THREE DECILES OF THE 1967 INCOME DISTRIBUTION
(Figure at top of box is mean change for group)

* Includes only families with same head interviewed in 1968 and 1969, head under 75 years old, and family income in the bottom 3 deciles of 1967 income (below \$4,649). The nine terminal groups explain 23.3 percent of the unexplained sum of squares in the original sample of 1,201 family units. Variables tried are listed below. Only those shown in the diagram above significantly reduced the unexplained sum of squares in the initial group.

Family size	Age of head	Time horizon	Risk avoidance
Sex of head	Background problems	County unemployment rate	Change in family work
Race of head	Current handicaps	Efficacy and planning	Change in family size
Employment status, 1968	Aspirations	Trust-hostility	

working in 1967. For families where this occurred the mean increase was \$2,913, while the average income change was only \$817 for those families where it did not. If, in addition, the husband scored high on our aspiration index, the mean increase in family income was \$4,431. It is, of course, quite possible that the wife's entry into the labor force was itself a response to the higher aspiration of the husband.

In the case of families in which a wife did not start to work, the interaction of head's age over 55 and low efficacy and planning score pushed the mean change in family income down to \$318. Even if the wife did not start to work in 1968, however, families with a male head under 55, who had completed high school and was working in 1968, had a substantial increase in income, \$2,230. Families headed by a female high school graduate, on the other hand, had even higher average increase in income, \$2,414, if the family also increased the amount of work members were doing.

In summary, it is apparent that short-run vari-

ation in poor family income is explained primarily by changes in labor force participation and by status variables such as age and education. Attitudinal variables come into play only after the others have done their work. It is probably true, however, that a family member's labor force participation is partially determined by the attitude of the head, and over a period of years change in poor family income may prove to be more strongly associated with attitudes. Further, it is interesting to note that background problems, current handicaps, and race did not come into play as predictors of poor family income changes.

Turning to changes in well-being measured by Orshansky score, we tried to explain them by the same AID algorithm and about the same set of predictor variables used for changes in money income. We did add a variable for change in family size, since the Orshansky measure is partially dependent on family needs. The results of the AID run and a list of the variables used are shown in Figure 3. As a frame of reference, the mean level of Orshansky score for all families in 1967 (in-

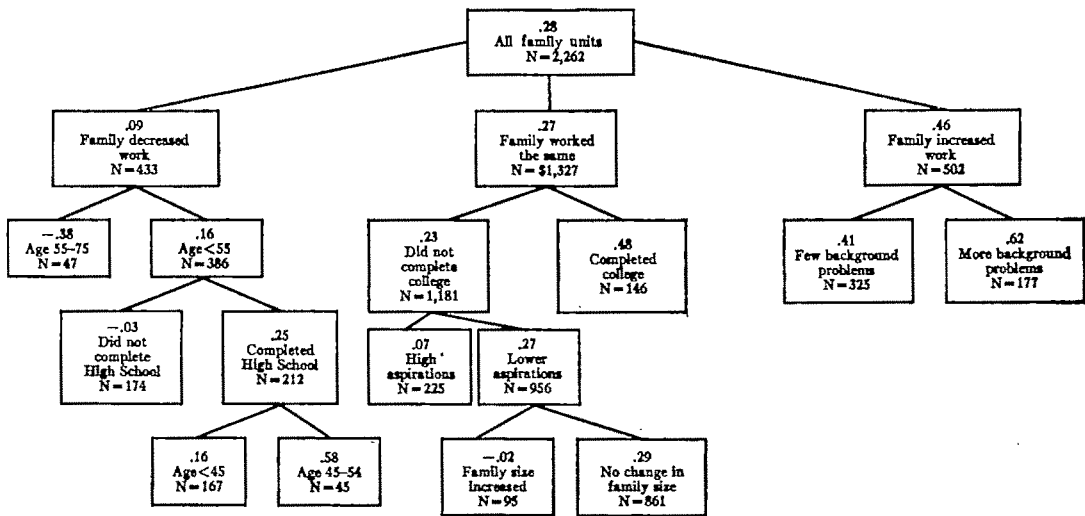


FIGURE 3

CHANGE IN ORSHANSKY SCORE OF FAMILIES* 1967 TO 1968
(Figure at top of box is mean change for group)

* Includes only families in which the head worked more than 1,000 hours in 1967 and 1968 and was interviewed in both 1968 and 1969. The ten terminal groups explain 6.7 percent of the initial sum of squares. The complete list of variables tried is listed below. Only those shown in the diagram significantly reduced the unexplained sum of squares in the original group.

Sex
Race
Disability
Education

Age
Background problems
Current handicaps
Aspirations

County unemployment rate
Efficacy
Trust-hostility
Change in family's size
Change in family's work

come year) was 2.6 and in 1968 (income year) about 2.8.

Orshansky score increased an average of .28 points for the whole sample of 2,262 families. The first and second splits came on changes in the families' work; that is, changes in the number of family workers and/or the number of hours they worked. Since the two splits occurred on the same variable, we have shown the three resultant groups across the top of the diagram.

Among families who decreased their work (left side of figure) age and education of head interact to produce five smaller groups which add to our ability to predict the change in Orshansky score. Families with heads 55 to 75 had an average change of $-.38$ Orshansky. This negative change may be attributable to men moving into partial or complete retirement and to the termination of work by older wives. Families with younger heads who had not completed high school did a bit better, but they too had lower Orshansky scores in 1968 than the year before, $-.03$ Orshansky. In these cases we suspect that their lower educational levels made them more susceptible to job loss, and their younger age is probably associated

with family growth and its negative impact on Orshansky score. Families headed by a man who had completed high school or went further and was 45-54 years old had a substantial increase in Orshansky score, $.58$, despite the reduction in family work. This group probably represents families in which the head has reached his peak earning years, older children are leaving home, and some wives are leaving the labor force in response to the high levels of their husband's income.

Of families with no change in work output between the two years, those headed by college graduates clearly did better, moving up nearly one-half Orshansky. But for families not headed by a college graduate our index of aspiration predicted the reverse of what was expected. Those with high aspirations had an increase of only $.07$ Orshansky, while the less aspiring group increased $.27$ Orshansky. Because this index has a fairly strong positive correlation with levels of income and need and a positive, although weak, correlation with changes in both income and Orshansky score, its inversion in this case is puzzling. Further exploration is clearly warranted, especially since the N's in the groups are large. This group of low

TABLE 5

DISTRIBUTION OF THE LOW-INCOME POPULATION* BY THE RELATIVE IMPORTANCE OF TRANSFER INCOME
IN 1967 AND 1968 FOR FAMILIES OF DIFFERENT TYPES

<i>All Families in Bottom 3 Deciles</i>			<i>Head Over 65 (30.6% of poor families)</i>		
	Less than 50% transfers 1968	More than 50% transfers 1968		Less than 50% transfers 1968 21.5%	More than 50% transfers 1968 78.4%
Less than 50% trans- fer income in 1967	48.2% $\bar{O} = 1.00$ $\Delta\bar{O} = .39$ (N=1094)	5.3% $\bar{O} = .87$ $\Delta\bar{O} = .05$ (N=101)	Less than 50% trans- fer income in 1967 (14)	10.4% $\bar{O} = 1.09$ $\Delta\bar{O} = .367$ (N=28)	3.5% $\bar{O} = .916$ $\Delta\bar{O} = -.02$ (N=9)
More than 50% trans- fer income in 1967	7.2% $\bar{O} = .81$ $\Delta\bar{O} = .76$ (N=114)	39.2% $\bar{O} = .80$ $\Delta\bar{O} = .11$ (N=611)	More than 50% trans- fer income in 1967 (86)	11.1% $\bar{O} = .89$ $\Delta\bar{O} = .87$ (N=28)	74.9% $\bar{O} = .79$ $\Delta\bar{O} = .11$ (N=190)
<i>Non-Aged Single Heads with Children (13.5% of poor families)</i>			<i>Non-Aged Single Heads—No Children (15.2% of poor families)</i>		
	Less than 50% transfers 1968 37.9%	More than 50% transfers 1968 62.1%		Less than 50% transfers 1968 57.7%	More than 50% transfers 1968 42.1%
Less than 50% trans- fer income in 1967 (41.4)	28.9% $\bar{O} = .92$ $\Delta\bar{O} = .34$ (N=176)	12.6% $\bar{O} = .87$ $\Delta\bar{O} = .20$ (N=48)	Less than 50% trans- fer income in 1967 (56.3)	49.1% $\bar{O} = .91$ $\Delta\bar{O} = .57$ (N=144)	7.1% $\bar{O} = .92$ $\Delta\bar{O} = .13$ (N=20)
More than 50% trans- fer income in 1967 (58.5)	9.0 $\bar{O} = .67$ $\Delta\bar{O} = .32$ (N=41)	49.5 $\bar{O} = .778$ $\Delta\bar{O} = .10$ (N=240)	More than 50% trans- fer income in 1967 (43.7)	8.6 $\bar{O} = .787$ $\Delta\bar{O} = .30$ (N=24)	35.0 $\bar{O} = .834$ $\Delta\bar{O} = .12$ (N=105)
<i>Non-Aged Married Heads with Children (31.8% of poor families)</i>			<i>Non-Aged Married Heads, No Children (8.4% of poor families)</i>		
	Less than 50% transfers 1968 87.7%	More than 50% transfers 1968 12.1%		Less than 50% transfers 1968 80.0%	More than 50% transfers 1968 20.0%
Less than 50% trans- fer income in 1967 (87.5)	84.6% $\bar{O} = 1.02$ $\Delta\bar{O} = .29$ (N=632)	2.8% $\bar{O} = .807$ $\Delta\bar{O} = -.08$ (N=18)	Less than 50% trans- fer income in 1967 (81.7)	75.2% $\bar{O} = .94$ $\Delta\bar{O} = .68$ (N=99)	5.5% $\bar{O} = .92$ $\Delta\bar{O} = -.22$ (N=6)
More than 50% trans- fer income in 1967 (12.4)	3.1% $\bar{O} = .76$ $\Delta\bar{O} = .90$ (N=17)	9.3% $\bar{O} = .705$ $\Delta\bar{O} = .14$ (N=72)	More than 50% trans- fer income in 1967 (18.2)	4.8% $\bar{O} = .75$ $\Delta\bar{O} = 1.21$ (N=5)	14.5% $\bar{O} = .97$ $\Delta\bar{O} = .01$ (N=26)

* Families in the bottom 3 deciles of distribution on Orshansky index in 1967. Families with a change in head between years are excluded.

\bar{O} = mean Orshansky score in 1967.

$\Delta\bar{O}$ = mean change in Orshansky from 1967-68.

aspirers was the only one for which changes in family size came into play.

Families which increased their work output (right side) had large gains in Orshansky score, but again an index, background problems in this case, worked in reverse. Use of this index in other analysis of the same data has indeed revealed a rather consistent reversal of the expected relation of background problems to change in well-being. This is particularly striking when the index is used to predict changes for a group which also has strong positive forces operating on it, such as heads with considerable education. It may be that something can be made of the idea that an uncomfortable past, coupled with an awareness that one can "make it," provides a considerable motivation toward economic success.

Comparison of the three AID algorithms we have discussed points up some interesting results. For changes in money income the situational variables, age and education, are of secondary importance after family size and attitudes. But when family size is partially taken into account in the Orshansky score these variables move up in importance, although they still follow after another variable change in family work. Family size may be acting as a proxy variable for amount of family work, in the money income AID tree, in which case there is no inconsistency. And the importance of changes in amount of family work is further supported by the run on low-income families which, after "employment of wife," looks much like the Orshansky score diagram.

The changes in work which come into play in the AID analyses primarily involve secondary sources of income such as the head's second job or the jobs of the wife or others in families where the head's main job is stable. A smaller number of these changes result in substitution of labor income for transfer income or vice versa. Table 5 focuses on changes of the latter sort for the low-income population.

Families in the bottom 30 percent of the distribution on Orshansky score in 1967 are categorized by the percentage of family income due to transfers in 1967 and 1968. Those in the lower left-hand cell of each group moved from more than 50 percent transfer income in 1967 to less than 50 percent in 1968 and will be termed the "start work" group. The "stop work" group in the upper right moved in the opposite direction. The distribution of these changes by family type is of substantial current interest.

The families of unmarried heads with children show the greatest shifting back and forth. Of these families, 9 percent "started work" while

12.6 percent "stopped work." Those who started work had a mean increase in Orshansky score of .32 over their mean level in 1967 of .67. This is a substantial relative increase, but, given their very low initial level, these families still fell short of the "poverty line" in 1968. Since the need standard was based on 1965 prices and was very spartan even then, these families can scarcely be said to have escaped poverty.

The net improvement in income position of these families came from a mean increase of \$2,056 in earned income which offset a mean loss of \$987 in transfer income. On the average, contributions of other adults in the family accounted for one-third of the increase in earned income.

The "stop work" group also appears to have increased their economic well-being. They had a mean increase of .20 Orshansky from an initial level of .87, placing them at a higher mean in 1968 than those who started work. This category, however, consists not only of those who actually stopped work but also of families whose transfer income increased with no decrease in earned income and of families where earners moved out but continued to help support the family via interpersonal transfers. Thus this group warrants further investigation before definite conclusions can be made.

The families with married heads who start work fit much more closely with the popular picture of getting a job and consequently getting out of poverty. The mean earned income of head and wife increased by \$3,865 for these families, easily offsetting a mean loss of \$1,029 in transfers and more than doubling their Orshansky score. These families, however, constituted only 3 percent of the married poor. About the same number stop work and suffer a net decline, ending up at a very low level of Orshansky score.

The aged poor constitute another idiosyncratic group. Their transfer income consists mainly of social security and pensions which are less likely to be "taxed" as earned income increases. The start work group among these families had a mean increase of \$1,500 from head and wife's income, supplemented by \$500 of other's income, while their mean loss of transfer income was under \$100. The large increase in Orshansky score for this group is less surprising in this light.

In overview, therefore, it appears that between the 1967 and 1968 income years a few families did move off welfare and onto the payroll and some of these families escaped poverty in the process. The percentage of the poor who did so was very small, however, and nearly as many moved in the opposite direction.

DISCUSSION

ARTHUR M. OKUN: The two papers presented at this session reflect the changing times. Income distribution—and especially the distribution at the low end—is back in the spotlight as an issue of social policy, and appropriately it is eliciting scientific attention. Early in the 1960's, concern with sharing the national economic pie was eclipsed by the urgent concern about the lagging growth of the pie. The profession was wise to focus on the aggregative issues in the early 1960's, and it is wise to shift focus onto the distributive issues as we enter the 1970's. Now that the nation has reaped the important benefits of economic growth and high employment, other problems deserve emphasis.

In the present stage of the study of the income distribution, part of the yield of research is in producing good questions as well as good answers. Both of the present papers make a contribution to knowledge and whet our appetites for more. They reflect different methodological approaches and stress different issues. They do share one substantive characteristic—pointing to the importance of labor force participation by the wife (or other secondary worker) as a key determinant of family income and income changes. Both papers leave us with some interesting conjectures about the wife's contribution to family income. Metcalf's¹ results suggest that the earnings potential of a woman may be especially sensitive to cyclical fluctuations—a proposition which can only be tested and verified if the wife's income is known separately from the rest of the family income. Smith and Morgan conjecture that the inferior performance of two-family persons in achieving income gains may reflect, to a large degree, their lack of opportunity for additional earners. Again the verification would require separate information on the components of family income.

In general, both papers remind the reader that family income is a complicated, summary statistic that reflects a lot of underlying determinants. As Smith and Morgan report, about 15 percent of the variance of annual change in family income comes from the wife's earnings, another 15 percent from changes in the labor income of other secondary earners, about 7 percent from changes in property income, and 45 percent from changes in the labor income of the family head. The last and largest component, in turn, compounds the impact of promotion, steadier employment, moonlighting, added amounts of overtime, and shifts of employment. It would be helpful to assess these various ways in which the labor income of the family head may change.

The Smith-Morgan findings on attitudinal variables make me want to know more about how attitudes affect income. I would expect attitudes to work by influencing family decisions on such diverse matters as how much to work, where to work, how to invest assets, whether to engage in job training, etc. But I do not know whether these are the channels the authors

consider important, because the paper does not develop their analytical views of how attitudes work.

Without such a model, there are a number of puzzles. First, it is not obvious why attitudes should have any predictive value on income change. As one would expect, the attitudinal variables correlate with concurrent levels of income. Presumably the aspirations, risk avoidance, and planning attitudes expressed by the family did not just arrive at the time of the initial survey. To the extent that a family has had stable attitudes for a long time, it is not clear that the attitudes should have any remaining influence on the change in income subsequent to the survey. Indeed, one might conjecture that 1967 attitudes work as predictors primarily by supplying insight into cases where 1967 incomes have an unusually large positive or negative transitory component. A family with a low 1967 income and with a set of attitudes normally associated with high-income families might be expected to have a higher 1968 income merely because its 1967 income was likely to be below its norm. One can imagine many good reasons for attitudes to influence subsequent incomes, but the authors' reasons for expecting such an influence are not clear.

We get some clues about the Smith-Morgan interpretation of attitudes from the kinds of questions that they ask in constructing their attitudinal indexes. People are virtually asked to forecast their incomes by some of the questions, e.g., do you plan to change jobs? Now it is entirely appropriate for a survey specialist to be interested pragmatically in any and all "noises" made by households that have predictive value. But whether he should want to call all such noises attitudes is another matter, and it may be more than merely semantic. One other clue the authors supply about their interpretation leaves me particularly puzzled. They say: "Attitudinal variables come into play only after the others [demographic and other "objective" family characteristics] have done their work." That implied sequential process is a mystery to me.

In his paper, Metcalf accomplishes a tour de force. He stretches the data beyond a reasonable breaking point and shows that they still do not break! He relies on two sets of unreliable data used in conjunction with each other. From what I know about the limitations of both the survey data and the national income data, I would have had great conviction that the kinds of manipulations that Metcalf performs would lead to nonsense results. It has no right to work! Yet it works remarkably well. If there are a few puzzling coefficients, the amazing thing is that they are so few, and that so many of the coefficients are sensible. This demonstration tempts one to become a true believer in econometrics, economic statistics, and economic theory all at once. As Metcalf states accurately in his own conclusions: "Despite major problems in coordinating . . . the data, a number of significant and plausible relationships between the size-distribution of personal income and changes in aggregate economic activity have been uncovered."

In the process, Metcalf belies an oft-expressed pre-

¹ Charles E. Metcalf, "The Size Distribution of Personal Income during the Business Cycle," published in the Sept., 1969, *A.E.R.* but presented at this session.

scription for empirical economic research which I have long regarded as a myth. The prescription goes: "Don't do anything fancy with crude data." It has always seemed equally plausible to me that one should do especially fancy things with unreliable data because it may take a highly sophisticated process of statistical refining to extract the valuable component of low-grade data.

Metcalf deals with three categories of families: those with a female head; those with a male head and with the wife in the labor force; and those with a male head where the wife is not in the labor force. Our recent experience of major increases in female participation reminds us that the shifts of the wife into and out of the labor force are themselves an important decision variable. In other words, the relative size of the two types of husband-wife categories should be made endogenous. In a footnote, Metcalf reassures us that he recognizes this piece of unfinished business.

It strikes me that regional cross-section data on family incomes may throw further light on some of the questions that Metcalf is asking. At any point in time, we do have some areas and localities of high unemployment and some of low unemployment within the nation; by studying income distribution within geographical regions, we may find differences attributable to varying levels of resource utilization. We need the help of regional data. After all, even a man with Metcalf's ingenuity cannot get more than one annual time series observation per year for the aggregate U.S. economy.

Finally, Metcalf's substantive results reinforce other important findings from Wisconsin's Institute for Research on Poverty that inflation does not harm the poor as a group. These facts should, once and for all, stem the temptation of the political economist to invoke the plight of the poor as one of the reasons for wanting price stability. There are compelling reasons to pursue price stability in our society, and inflation may well be, as the cliché goes, the cruelest tax. But inflation is cruel because of horizontal (not vertical) inequity. Indeed, it seems to have a haphazard impact among families within the same income group which outweighs its redistributive impact among groups. For example, welfare recipients and the aged poor may be hurt by inflation even though the predominant effect on the poor is that of improved job opportunities. Inflation is an inefficient and inequitable poverty program; we can certainly design better ones at lower overall cost to the nation. Certainly the Wisconsin results do not justify inflation. But they should sharpen our efforts to define and assess the welfare costs of inflation. And they should remind us that, when the nation is deliberately pursuing a policy of economic slowdown in order to achieve price stability, it has a special obligation to intensify other efforts to relieve the plight of the poor.

WILLIAM H. BRANSON: The Smith-Morgan and Metcalf papers raise many questions—perhaps, on balance, more than they answer—about two different aspects of the income distribution problem. The first issue concerns the distribution of permanent incomes.

What causes this distribution to be what it is, and what policies should be aimed at improving the permanent incomes of people in the (however defined) lower tail of the income distribution?

The other issue concerns variability of income. What proportion of the people in the lower end of the distribution in any given year are there because of a temporary drop in income? Federal government policy toward the latter kind of measured poverty should probably run along the lines of unemployment insurance, while policy toward permanent income poverty should probably involve a combination of a large-scale permanent transfer program, such as the Family Assistance Plan (FAP) with a decently higher floor, and education and training programs.

The paper by Smith and Morgan focuses on the variability of income, measuring, in a 4,500-family sample, the percentages of families who changed income deciles from 1967 to 1968. Movements in income are measured both in absolute terms and relative to the families' poverty level income as given by the Social Security Administration. These changes in measured income are then related to a number of indexes of social attitudes. The basic findings of the study, to date, seem to be that (a) income is surprisingly variable, although no summary statistic of variance is presented, and (b) families with the good middle-class virtues of high aspirations, self-images as careful planners, and risk avoidance did the best on changes in income from 1967 to 1968.

There are two basic questions I would hope the Smith-Morgan study will answer fairly soon. First, to what extent in a given year is the population of some specified lower tail of the income distribution, say that below half the family median income, due there to low permanent income as opposed to temporarily low income? Second, how does the variance of income vary by income decile? This information would be useful in designing income maintenance experiments.

Since the Smith-Morgan paper is essentially a progress report on a continuing project, it is only natural that it leaves a few loose ends still dangling. I thought it would be useful to them to point out some of these questions.

1. A good deal of space is given to measuring incomes relative to the SSA poverty line. This seems to me pretty irrelevant. Rather than spreading the use of arbitrary poverty lines, it might be better to put incomes on a per capita basis with some simple allowance for economies of scale.

2. After finding lots of downward shifting in the high income deciles, the authors point out that this may be because when families split up between 1967 and 1968 the entire family income in 1967 was imputed to each of the 1968 units. Why? How much does this affect the results?

3. When income variability is measured using income relative to the SSA poverty income, the authors say variability is higher than when measured on absolute income levels. This seems to imply a negative covariance between actual income and poverty-level income (as family size changes) across families. Yet

later a positive covariance is reported. How are these findings reconciled?

4. In Table 3 the authors give a negative correlation coefficient of .22 between changes in earned income and changes in transfer income. The sample includes families with age of head up to 60, so what proportion of this effect is due to retirement?

5. Concerning the social determinants of income change, what is the reader to infer about causation? This is important for understanding the poverty problem and forming policy. For example, what do we make of "... high aspirations and a college education interacted to yield the largest mean increase in income, \$2,312, among large families headed by risk avoiders." Does education and high income lead to high aspirations? Or perhaps high aspirations and high income lead to a high consumption of education.

6. There seems to be some confusion between things affecting income levels as opposed to income variance. Things like sex, race, and disability probably affect permanent income levels more than income variability.

7. Finally, I must express skepticism about the social indexes. For example, people with preferences for saving over spending get points for efficacy and planning. But the life cycle hypothesis tells us that the young, and in general people with current income less than permanent income, should be borrowing if they are good planners. In another case, "not employed during year" adds to the aspirations index score. These points in the appendix need some discussion, anyway.

The Metcalf paper relates movements in the size distribution of income to various macroeconomic variables. This work could make a useful contribution to macroeconomic models both by giving us a clue on the effects of various government macropolicy changes on the income distribution and by translating these into effects on economic activity through, for example, differing marginal consumption propensities by income class.

To measure income distribution Metcalf uses the income cutoffs for the top and bottom deciles of a displaced lognormal distribution, relative to the median, for three family types: intact families (both husband and wife present), with wives in and out of the labor force, and families headed by women. This measure of distribution looks sensible, as far as I can tell.

The only general problem I have with the paper is that in the equations relating the upper and lower decile incomes to things like wage rates, unemployment, participations rates, etc., at several points it is very hard to see how the effects of these variables get disentangled. One result is that seemingly odd results get reported without much explanation. Perhaps this is due to squeezing a larger paper down for publication.

There are two examples of this confusion, at least in this reader's mind:

1. The corporate profit share of GNP appears only in the equation for the upper decile income of the distribution for intact families with wives in the labor force, and with a negative coefficient. This mysterious

effect is explained by the suggestion that "... the corporate share of GNP, except for the dividend component, represents the portion of GNP that is retained from personal income. While upper income families ultimately have an interest in the increased corporate share, in the short run their share of measured personal income declines." But if this were the explanation, the variable should also appear in the upper decile equation for intact families with wives not in the labor force. And besides, the income distribution data in the study are from the Census' Current Population Surveys, not National Income Accounts personal income, so this kind of statistical relation between family income and corporate profits will not necessarily hold.

2. Another mysterious variable is the price level in the same equation. In an equation that includes the employment rate, the participation rate, and real wage rate, the price level has a positive sign. This is explained as "... probably attributable to the labor-force-oriented character of the group." Given the presence of the other variables, the only explanation I can develop is that in inflation the salaries of high-income intact families with wives in the labor force rise relative to the median income of such families. But perhaps the profit share and price variables are just canceling each other out, since they tend to move together over the cycle.

Perhaps, in general, Metcalf should see if some comprehensive cyclical measure, such as the concurrent indicators in the *Business Conditions Digest*, gives results that are consistent with those he reports, but statistically worse. As the paper stands now, it is very hard to sort out effects of the macro variables in some of the equations.

MICHAEL PIORR: Among the problems aggravating the society and disturbing our campuses and cities, income distribution is the one to which economic research seems most directly relevant. The problems of income distribution are defined within the framework of traditional economic analysis: they appear susceptible to treatment with conventional tools and they have received relatively little attention in the last thirty years so that the potential return to work in this area would seem high. Until we find a tractable definition of imperialism, therefore, these problems are about as close as economics can come to being where it is at. And that makes these papers exciting and interesting to deal with.

For me, anyway, it also makes them difficult to criticize. I found no want of points in the papers with which to take issue. But I felt that much of what I had to say was carping. Given the importance of the field and the paucity of work in it, most of my points of criticism seemed at most a question of priorities and, possibly, only a question of taste. Nonetheless, in trying to sort out what these papers had done from what I would have liked them to have done, I was not all that persuaded that the authors' priorities and tastes in this matter were better than my own. This raises the question of what standards to use in examining work in this field, and I would like to use part of my time to talk briefly to that question.

It seems to me that an assessment of work on income distribution requires both normative and positive standards. The normative requirements are for ways of measuring and displaying the distribution of income which separate those disparities with which the society is concerned from those which are more acceptable or, at least, of lesser interest. Economists have grown lazy in making normative assessments of this kind. We have been able to rely for the last twenty-five years on the standards set forth by the Employment Act of 1946. But the new policy goals require, I think, a more profound appreciation of what is motivating ourselves and our clients. Obviously, one could speak at length to this problem. I will hold myself to three, I trust, acceptable points. First, an income cutoff point defined in absolute terms has proved a relatively unsatisfactory measure of the society's concern. Second, of considerable concern are the racial dimensions of the distribution. Third, of considerably less concern are distributional dimensions generated by stage in the life cycle. And that for these reasons racial and life cycle dimensions need to be separately identified.

The positive dimensions of income distribution problems are relatively straightforward. We want to know how and why the economy generates given distributions and how these are related to others of our economic and social goals. Because interest in income distribution is at root a product of specific policy concerns, two additional criteria might be added. First we need to understand, not just any aspects of the distribution which catch our intellectual fancy, but those aspects which are of normative interest. Second, the positive understandings should ideally point to policy instruments which enable the society to change the distribution in a manner more conducive to social welfare.

Given these standards of judgment, it is difficult to fault the Metcalf paper. The problem which he has chosen to work on is to fit the income distribution into econometric models of the American economy. A great deal of what we know about the way the economy behaves is summarized in such models, and there is probably no greater single contribution one can make toward a positive understanding of the income distribution than this. The limitations of the analysis are inherent in the problems he has chosen to work with. The particular distributions with which he is working are not of greatest normative interest, but they are the available annual data. Most econometric models have proved, I think, a relatively poor source of new policy instruments, and it is my own feeling that we will need a much more complex understanding of the microeconomic processes which generate such models before we can effectively change the distribution of income in a manner consistent with other social and economic goals; but so much is invested in this approach to understanding the economy, that it would be hard to question its selection as a starting point for further research. Moreover, neither the normative nor the positive failings of Metcalf's approach

are irremediable—and he has indeed promised to remedy some.

I find the Smith-Morgan paper less satisfactory in these terms. They are collecting their own disaggregate data and have the option of putting it together in any way they choose, and while what they have done with it is interesting, I wish they had put it together in a somewhat different way. The aggregate income distribution with which they are working is, I think, one of the least important aspects of the distribution either positively or normatively. I wish they had devoted more explicit attention to separating life cycle changes from other influences upon the distribution. This is one of the greatest potential contributions of a panel survey. It is, as I have suggested, of major normative interest. And its positive interest is obvious. From a normative standpoint, at least, the racial distribution of income and the distribution of change is also of obvious interest. I would have liked to see more experimentation with the distribution within groups defined in various different functional ways: by source of income, by days worked during the year, by type of occupation, etc. As the number of years covered by the survey expands, the criteria available for separating out groups whose distribution is influenced by similar factors will increase. The panel survey will also provide bench-mark data for linking the sources used by Metcalf to series of greater positive and normative relevance; these opportunities also will be lost if the data are not displayed on a more disaggregate level.

Finally, I was disturbed by three technical problems in the Smith-Morgan paper. First, it seems to me that several of the attitudinal indices may be reflecting income level rather than attitudes. Thus, for example, the risk avoidance index includes items such as car ownership, car insurance, and medical insurance and the low aspirations index scale may simply reflect the achievement of aspirations permitted by income.

Second, the use of absolute income change rather than percentage income change is debatable. If the underlying economic processes occur in percentage rather than absolute terms, change in income will tend to be affected by the level of income. If the value of attitudinal indices is a product of income level and income change is also related to income level, this may produce high correlations between attitudinal indices and income change which have little real economic meaning. The use of absolute, as opposed to percentage, change in income as the dependent variable may be responsible for the greater success in explaining the first three deciles of the income distribution than in explaining the whole thing. Third, most of the demographic and economic variables used to explain income change are levels or states; I think that changes in income might be better explained by changes in these variables.

Finally, let me say again, that I find the research discussed in both of these papers extremely interesting and I appreciate the opportunity to review and comment on them while they are still in progress.

MARKET AND PRICE MECHANISM IN SOCIALIST COUNTRIES

THE RISE AND FALL OF ECONOMIC REFORM IN CZECHOSLOVAKIA*

By OLDŘICH KÝN

Charles University, Prague, and University of California, Berkeley

Before August 21, 1968, the Czechoslovak economic reform had been considered the most far-reaching reform in Eastern Europe. It is not anymore. It still may be interesting to review briefly the short period of reformist movement in Czechoslovakia and show how far the economic reform really went and at what point the process was reversed.

To measure the progress of the economic reform, however, is not easy. The result is very different, depending on whether we base our judgment on: (a) writings of theoretical economists, (b) writings or statements of planners or party leaders, (c) documents already approved by government and party organs, or (d) the actual state of the economy.

Passing from (a) to (d) our impression might be that less and less was actually accomplished. Some people might, therefore, be inclined to consider the real state of the economy as the only relevant question. It is, however, not so. In a command economy it is not only the day-to-day coordination of economic processes but also the process of organization that is highly centralized. Any institutional changes have to be decided at the top and approved by the Communist Party Central Committee. Therefore, the shift in the ideas and preferences of the ruling elite was extremely important if the economic reform was to be implemented. Without their approval no reform could have taken place nor could a proposal for reform be publicized. They would never have approved anything like the market, free pricing, or workers councils if their views had not been drastically changed. To develop the new theory and persuade people in the ministries, the Planning Commission, and the CP Central Committee to accept it was not at all easy. But once it was done, a substantial step toward reform was guaranteed.

During most of the Stalin era, and even for some time after his death, the development of economic theory did not have any important influence on economic policy. The rulers would

first make a pragmatic decision; then the role of theory was to find an *ex post* theoretical justification of it.

But this time it was different in Czechoslovakia. Theoreticians first made a proposal to change the economic system, and then—surprisingly enough—politicians did accept it, though not immediately and without hesitation. This can be partly explained by the more liberal attitude towards economic science in the early 1960's. Different views were allowed to be expressed and discussed. With better tools and less danger of the accusation of revisionism, economists were encouraged to suggest new ways to organize the economic system. But probably more important was the fact that Novotny had no other solution at hand. The economic situation in 1962–63 was so bad that he clearly saw that something had to be done with the economy.

So it happened that people from research institutes and universities rather than planners and party "aparatchiki" originated and stimulated the movement towards the economic reform, and thus played a true leading role in the Czechoslovak economy for a moment. And that was a sin which must not be forgiven. "The more serious shortcoming was, that from the early beginning of the work [on economic reform] a group of self-confident economists was formed, which in fact monopolized all further work on the reform, and gradually appropriated the right for issuing final judgments."¹

I. Theory

During the 1950's Czechoslovak economics belonged to the most sterile and dogmatic in Eastern Europe. Unlike Poland, the new wind came very late. In 1962–63, Czechoslovak economists were still discussing whether mathematical methods can in any sense contribute to the development of economic theory, and so-called "bourgeois" economic theory was virtually unknown. Since that time, however, very much has happened. Especially, the young generation of econo-

* I am indebted to Professor Howard Sherman for his help and valuable comments.

¹ Dr. V. Čermák, ing. H. Kyálka in *Hospodářské noviny*, No. 46, Nov. 14, 1969, p. 3.

mists quickly learned how to use the modern tools of economic analysis, and even the older generation abandoned most of the ideological prejudices which had hampered the previous progress in theory. Still, much remained to be desired. The acquaintance with modern analytical tools was casual, as it probably could not have been otherwise. Thus, for example, the basic ideas of the reform were originally developed without any knowledge of the famous Lange-Lerner model of market socialism. Moreover, during years of discussing the advantages and disadvantages of the market mechanism, Czechoslovak economists did not realize the importance or even the existence of "externalities."

These are the main areas in which considerable change in economic thinking has taken place:

1. *The Role of Ownership and Economic Laws.* Dogmatic Marxist theory interpreted ownership (private under capitalism; state and collective under socialism) as a "substance" which determines the nature and character of economic systems and, at the same time, determines economic laws which "rule" the economy. Thus private ownership was conceived as a final cause of exploitation and anarchy under capitalism, while socialist state ownership was supposed to be a necessary and sufficient condition for the working of "the basic economic law of socialism" and "the law of planned and proportional development."

At the beginning of the 1960's some Czechoslovak economists tried to return to Marx's notion that legal ownership does not cause economic behavior but rather is a reflection of basic economic relationships, as for example cooperation and division of labor, control over the process of production and forms of distribution. The new approach shifted their interest from speculation about "substances" and "laws" to the analysis of the actual behavior of the agents of economic systems. Eventually, most of the metaphysical "economic laws" of Stalinist economics quietly disappeared from economic theory.

2. *Marxist and Non-Marxist Economics.* Gone are the days when only the officially approved (supposedly Marxist) views were considered to be "scientific" and everything else was "bourgeois" or "revisionist" and invented only for "fooling the working class." Recently the apologetic character and sterility of dogmatic Marxism have been criticized. It is now said that modern Marxist economics ought to absorb all positive methods and theories (after they have been cleansed of all antisocialist ideological interpretations) and ought not be limited by concepts and tools developed in the nineteenth century. For some, it has

even become irrelevant whether their views are called Marxist.

3. *Character of Socialism.* Referring to Marx's *Economic and Philosophic Manuscripts*, many Czechoslovak economists asserted that the essence of socialism is not only the abolition of private ownership and exploitation, but first of all the humanization of society, abolition of alienation, and true liberation of man from passive dependence on nature and social institutions. Thus, planning and central direction were conceived only as the means and not the end in itself. It was understood that socialism is not a society with ideal harmony, where everything goes as in a "well-oiled machine." The obvious facts of stratification and conflicts of interests were theoretically accepted. Some even lost their belief that Czechoslovakia as well as the Soviet Union and other East European countries are really "socialist."

4. *Critique of the Command Economy.* The command system was criticized for inefficiency and waste of scarce resources, permanent disequilibria, extensive growth with declining productivity of factors (both labor and capital), causing deceleration of the rate of growth, fluctuations in the rate of growth, misallocation of investments and distortion of structure (the steel fetish), insufficient growth of the standard of living, rigidity, hampering innovation, administrative monopolization, informational inefficiency and distortion of information, inefficient foreign trade, lack of long-term planning, autocracy, etc.

5. *Market Socialism.* Market socialism was considered to be a much better system than the command economy. The aim, however, was not the "pure market system" but rather a combination of central planning with spontaneous market forces. The market ought to provide both stimuli and information for day-to-day coordination of most economic processes, while central planning ought to guide development towards given social ends. The conditions necessary for making market and planning compatible and also the peculiarities of the market mechanism in the absence of private ownership were discussed.

6. *Price Theory.* Price theory emerged as a combination of Marxist (labor theory of value) and non-Marxist (marginal utility) price theory. "Rational types of prices" were also discussed, but unlike theories in the U.S.S.R. and most of the East European countries, the need for free pricing (which would equilibrate supply and demand) was postulated. Free pricing refers here to the situation where only a small part of all prices are directly fixed by the government, while

the remaining prices are fixed by the firm, though they are under central price controls.

7. *The Real Autonomy of the Firm.* The real autonomy of the firm in decision making was considered as crucial for economic efficiency. This implies that planned targets should be given to firms only as "orientation" and not as an order which must be fulfilled at any costs. Firms should be free to look for new combinations of productive factors and to produce any product they find profitable. "Socialist entrepreneurship" was conceived as necessary for promotion of innovations and thus for keeping pace with technical progress in the industrially developed countries.

It would be possible to enumerate many more aspects of the new stream in economic theorizing in Czechoslovakia. Very interesting, for example, was the rehabilitation of the concept of spontaneity, which was not considered as an evil anymore. Conscious central decisions need not necessarily be rational or optimal; and, therefore, there is no reason to believe that suppression of spontaneity must be good in itself. Reference to Darwin's law of natural selection is relevant here. It shows that such complex systems as the human brain could arise as a result of spontaneous processes.

Let us step down from the realm of pure science to the views of the very people who make the decisions in the economy. To be brief, we shall call them simply planners. Their first reaction to the proposal of theoreticians was negative, as could have been expected. In the fall of 1964 the commission consisting of both theoreticians and planners was established for drafting "The Principles of New System of Management." This first official document (approved by the Central Committee in January, 1965) was in many respects revolutionary, but it still contained statements about the superiority of plan over market. According to views prevailing among planners at that time, the central plan must determine rates of growth and all important "proportions" in the economy, while the market must be subordinated to the plan and be used only as an "instrument" for implementation of the plan. The opponents of this view did not dispute the usefulness or even the necessity of central planning for achieving smooth and optimal growth, but they did consider the market as something more important than simply an instrument for plan implementation. Market structure contains channels transmitting information on consumer preferences and scarcity of resources which in the command economy is difficult to obtain. Apart from decentralized coordination, the market mechanism generates

parameters (prices) necessary for optimal planning. The remarkable progress in the views of planners is seen from this statement of the deputy chairman of the State Planning Commission in Czechoslovakia, Miloslav Kohoutek:

The plan ought to respect the market as a genuine economic category. At the same time, however, it is bound to be an act of autonomous decision-making, that should affect economic processes and also the market in a sense. Therefore, those concepts, which regard the market as some sort of an instrument for carrying out the plan, are unrealistic, since the market cannot be considered as a category subordinated to the plan. There exists a definition of this problem, which is used in Western countries now and then and which, in my opinion, quite adequately regards the plan as some kind of a model of the future market. It reveals, though not fully, what should be embodied in the plan no matter which society is using it, as long as there exists the market.³

II. The Reform

Now we shall survey briefly the changes in the economic system which had been already implemented (or at least approved) before the policy changed to anti-reform.

1. *Planning.* At the beginning of 1967 the so-called "directive" character of planning was abolished. The targets (with some, but not very many, exceptions) given to enterprises were not orders, but rather suggestions what to produce. In 1968 and 1969 annual central plans did not exist. Instead of plans the government issued so-called "Economic Guidelines" containing parameters for economic policy and information about desirable output for whole branches of the economy. In preparation of its own annual plan each enterprise was to take into consideration the Economic Guidelines of the government but was not given any order as to how much to produce. The sum of the plans of enterprises was not necessarily equal to the targets given for the whole economy by central planners. In 1969 the enterprises planned in ways that deviated from the economic guidelines of the government as follows: (a) a lower rate of growth of output (4.6 percent by the sum of firm decisions compared with 6 percent in the central guidelines); (b) lower export to and higher import from the West (with an imbalance of 3 billion Kčs); (c) higher demand (500 million Kčs more than the guidelines) and lower supply (7 billion Kčs less) of investment goods; (d) higher employment (increase in 1969 to be 41 thousand while the actual increase in 1968 was only 4 thousand); (e) higher wages; etc.

³ *Czechoslovak Economic Papers*, No. 10, 1968, p. 126.

2. *Prices.* The price reform which took place in January, 1967, increased the level of wholesale prices by 30 percent (compared with an expected 24 percent), and also redistributed a considerable part of national income from central funds to profits of enterprises (about 16 billion Kčs more than expected). The objective of price reform was to create conditions for: (a) introduction of a uniform capital charge and a uniform rate of taxation (which was achieved); (b) substantial reduction of subsidies (only partly achieved); (c) shift from centrally fixed to free and controlled prices (only partly achieved). The situation of prices was as follows:

tiated "deductions from profit" or "transfers" to the state budget, which depended on fulfillment of plan indicators, uniform taxation was introduced. Taxes consist of a 6 percent capital charge, 18 percent of "gross income," and the so-called "stabilization tax."

5. *Foreign Trade.* The process of gradual liquidation of the "state monopoly of foreign trade" started. In some branches competing trading organizations were created and some big industrial firms were allowed to sell or buy abroad. Most prices of imported and exported goods were in the category of "free" prices. The system of differentiated surtaxes or subsidies to the export

CATEGORIES OF PRICES
(In percent of total sales)

	At the End of 1968			Expected at the End of 1969*		
	Fixed	Controlled	Free	Fixed	Controlled	Free
<i>Wholesale</i>	16	80	4	15	40	45
<i>Retail: Total</i>	77	—	23	45	16	39
Food.....	72	—	28	63	—	36
Industrial goods..	82	—	18	25	33	42

* Official estimation.

"Controlled prices" are those prices for which the firm must stay within the maximum increase allowed by the government for the aggregate price index of all products of given enterprise but may vary at will the prices of each individual good within the category. The idea, of course, was to allow flexibility while preventing inflation. This category has previously been misinterpreted in the West to mean simple maximum and minimum limits for individual prices.

3. *Decentralization of Investments.* Investments in industry were financed basically from retained profits and bank loans. Investment subsidies from the state budget were considerably cut. The total amount of investment (including investments in the "unproductive sphere") increased from 43 billion Kčs in 1965 to 70 billion Kčs in 1967, but investment subsidies from the state budget to so-called "economic organizations" fell from 19 billion Kčs in 1965 to approximately 10 billion Kčs in the years 1967-69. The investment credits and loans issued by the State Bank amounted to 13 billion Kčs in 1965 and 25 billion in 1967, because of the shift from central to decentralized financing.

4. *Taxation.* Instead of the system of differen-

and import of individual products was to be abolished in five years (20 percent reduction each year). It was believed that this would improve the efficiency of foreign trade and eventually lead to convertibility of the crown.

6. *Autonomy of Firms.* At the beginning of the economic reform all enterprises were organized in about a hundred associations or trusts. The intention was to decentralize the controls from ministries to associations, but the ministerial system was in fact not abolished as was originally expected. Since they acted as pure monopolies in the market, creation of associations was in no sense good for promotion of competition, which was so badly needed for the functioning of the market mechanism.

In 1968 it was already clear that some serious steps had to be taken to guarantee the real autonomy of the firms. First outlines of the "enterprise bill" containing the idea of "enterprise (workers) councils" began to be shaped. At the beginning of 1969 the bill was finally drafted and was expected to be approved before long by the National Assembly. According to the bill, enterprises would be governed both by professional managers and enterprise councils. The idea was

to combine the skill of professional managers with the democratic principle of self-government. Managers would have the sole responsibility for current management (without any intervention by the councils) and for preparation of long-term projections for enterprise (requiring approval by the councils). Councils, as representatives of the owners, would elect directors for a four- to five-year period and would decide about distribution of net income.

The bill distinguished three different types of enterprises according to ownership:

a) "Public enterprises" (about 90 percent of all enterprises), with councils elected by all employees of the enterprise, would be in collective ownership. Four-fifths of the council should be enterprise employees and one-fifth should be external members (possibly representatives of government, the state bank, experts, etc.).

b) "State enterprises" (such as the state bank, state railways, etc.) would remain in state ownership. The larger part of their councils would be nominated by the government.

c) Enterprises established by other enterprises or by state organs. These enterprises would be owned by their founders, who would nominate

by M. Barta in the journal, *Politická ekonomie*, No. 8 (1969). The first striking point is the composition of the councils. Even when they are elected in free democratic elections by all employees, 73 percent of the members are engineers, 22 percent workers, and only 5 percent administrative workers. Over 80 percent were employed in the enterprise for more than ten years, and most of them were highly skilled. The average level of education of the members of councils (29 percent with university education) was probably higher than the average level of education of the directors of these enterprises. M. Barta concludes: "This finding suggests that enterprise councils are the true industrial elite . . . creation of councils can be considered as a protest of workers against the dilettantism of bureaucrats. . . ."

Another interesting finding from the inquiry is that only 5.9 percent of the chairmen of councils answered that the main role of the council should be the maximization of income for the workers of the enterprise, while 31.6 percent thought that councils should mainly decide about the long-term development of the firm. The following table compares the views of directors and chairmen of councils.

THE ROLE OF COUNCILS

	According to the Views of					
	Chairmen of Councils			Directors		
	Decide	Approve	Consult	Decide	Approve	Consult
Long-term economic policy . .	51.1	38.0	10.9	23.2	62.2	14.6
Annual plans	3.2	53.2	43.5	2.4	36.6	57.3
Rules for wage and bonus determining	—	36.6	62.3	1.2	27.5	62.5
Distribution of income	16.3	76.1	7.6	10.8	68.7	19.3

most members of their councils.

During late 1968 councils were established in some enterprises. In January, 1969, there were councils already in more than 115 enterprises (most of them in industry), and creation of councils continued in the first half of 1969. It is worth mentioning that there are no more than 2,000 enterprises in the Czechoslovak economy in total, and only about 700 in industry.

Interesting information about 95 councils comes from an inquiry carried on in January, 1969. Some results of this inquiry were published

III. The Anti-reform

The political situation had substantially changed after August 21, 1968, but it seemed for a while that the economic reform would continue in the same direction as before. There were severe attacks on Šik's theory from the Soviet Union, but a critique of his economic theory could be considered as nothing more than a pretext for condemnation of his political behavior.

There were not many reasons to believe that Soviet leaders strongly opposed any basic aspect of the economic reform—except possibly the idea

of workers councils and the loosening of the direct control of the Communist Party over the economy. The Soviet arguments concerning foreign trade were not convincing. It was clear to anybody who knew the basic facts that for Czechoslovakia, due to strong competition and the low quality of Czechoslovak goods, it would be very difficult to increase exports to the capitalist market. There were no indications that the steps towards liberalization of foreign trade which had already taken place would lead to any considerable shifts in foreign trade. Furthermore, the economic reform in Hungary, very similar to the Czechoslovak reform, has continued without attracting the attention of the Soviet Union.

Thus there existed a hope that the end of Dubček's political system need not necessarily mean the end of the economic reform. It is true that the situation in the economy was not favorable to further decentralization. But in spite of shortages, disequilibria, and inflationary pressures (created or strengthened by the presence of large army and military operations on Czechoslovak territory), binding annual planning was not restored, the share of free prices was expected to increase, more and more workers councils were established, and a legal guarantee for the autonomy of firms was still in preparation.

As late as May, 1969, a group of economists from the Economic Institute of the Czechoslovak Academy of Sciences and from the secretariat of the government submitted "An Outline of the Conception of the Further Development of the Economic Reform." This outline took into account the very difficult situation in the economy and was therefore aware of the necessity of adopting some temporary control measures, but it did not make any substantial concessions as far as basic principles of the reform were concerned. This proposal was rejected. From that time, voices criticizing Šik's reform began to appear.

On May 23, Jaroslav Jirásek published in *Hospodářské noviny*, "An Outline of the Ideology of the Economic Reform in CSSR," in which he criticized the "extremism" and "revisionism" of 1968 and stated that it is impossible to continue the economic reform in its original form. According to his opinion, the economic reform neglected the importance of professional management and central control, overestimated the role of market self-regulation, and falsely put the dilemma of "the plan versus the market" as the central problem of reform: "This dilemma is in fact no dilemma, because the centralist administrative model is so discredited that to choose this model would mean to reverse the course of history . . .

there is no other way than the market character of the economy. . . ." He proposed the reorientation of the reformist movement towards the increase in the skill and authority of economic management. He suggested that the reform should proceed in three stages: (1) In the first stage—approximately 1970—he suggested the reintroduction of the administrative control of prices and wages, cutting investments in order to reduce overemployment, and finally the restoring of "the chain of power." (2) In the second stage—approximately 1971–74—to use the authority of the state and the management elite to achieve economic growth. (3) In the third stage—1974–80—"full deployment of the advanced economic system, based both on planning and the market mechanism."

In the meantime the economic situation has considerably deteriorated. In the middle of 1969 industrial and agricultural production stagnated, while money incomes continued to climb at an unprecedented rate of growth in Czechoslovakia, followed closely by retail prices. Shortages of almost all kinds of consumer goods, with resulting long queues—not seen for several years—began to be common in the market.

In this situation several practical measures were adopted:

1. *Planning.* First steps were taken to restore the binding character of the central plan. For the current year the system of "agreements" between government and general directories was established. These agreements were used as a device which would force enterprises to conform with central directives, particularly to produce more and to slow down the growth of wages. For 1970 restoration of annual planning is being prepared.

2. *Prices.* From July 1, 1969, most of the free prices were transferred into controlled prices, and maximum limits for these prices were established. Introduction of price ceilings was supposed to be only a temporary measure. The idea of central regulation of prices through a system of "price agreements" was not rejected.

3. *The Autonomy of Firms.* The adoption of the "enterprise bill" was postponed and the role of workers councils was curbed. "It is necessary to strengthen the position and authority of directors in enterprises. . . . We suggest screening of all existing workers councils . . . we can approve only those workers councils which had been established before April 1, 1969 . . . under no conditions can councils created after May 15 of this year be approved."

¹ Czech Minister of Industry F. Čihák in, *Hospodářské noviny*, No. 26, June 27, 1969.

4. *Taxes.* The tax reform changing the taxation of "gross income" into separate taxation of profit and wages was in preparation already in 1968, so it cannot be conceived of as an anti-reformist act.

In the second part of the year the critique of Šik's reform was amplified, and the necessity of reinforcing the role of central planning more frequently postulated. For example, the federal Minister of Planning, F. Vlasák, wrote together with L. Říha at the beginning of September: "We must increase the authority of planning. . . . We shall have to take into account the fact that Czechoslovakia is an integral part of the socialist commonwealth. . . . It is unrealistic to assume that some basic aspects of systems of economic control could substantially deviate from the principles applied in the other countries. . . . It will be necessary to return to some of the rational principles of the economic system that were developed in the years 1956-58."⁴

But even this kind of flexibility did not save him from being purged, together with some other economic ministers who were too closely associ-

ated with the economic reform.

The new federal Minister of Planning, Václav Hůla, took the new and more austere standpoint. He argues that the original proposal of reform accepted by the Communist Party Central Committee in 1965 was correct because it unambiguously accented the improvement of planning. "Already at that time . . . revisionist tendencies, opinions, and attitudes began to appear in the field of theoretical economics. They were hiding under the slogan of deepening and speeding of reform. Because they were not identified as revisionist and they were not overcome, they finally created a consistent rightist revisionist platform. . . . After January, 1968, the open antisocialist interests of reaction entered into the play. . . ." The only way out is "to restore the leading role of the party in the economy."⁵

At this moment a commission is still working on the new conception of economic reform. What will come from it is not really clear. Much of Šik's reform has already been scrapped, but not everything. The new leaders at least still insist that they do not want to restore the administrative system of planning.

⁴ Doba, Sept. 4, 1969.

⁵ *Hospodářské noviny*, No. 48, Nov. 28, 1969.

ECONOMIC REFORMS IN HUNGARY*

By RICHARD D. PORTES

Princeton University

Some knowledge of the background underlying the Hungarian economic reforms of 1968 is essential to any consideration of their content and prospects. First, one must recall that partial decentralization was attempted a decade ago. Because the measures of 1957 significantly increased the flexibility of the system but failed to change its fundamental characteristics, increased pressure on the economy in 1959-60 brought uncoordinated, piecemeal recentralization. The lessons were clear to the reformers of the mid-1960's, who had also benefited from experience with the new techniques of economic management introduced in 1957.

Second, the reforms were carefully prepared. The Central Committee decided in December, 1964, to initiate full-scale discussions, but no significant changes were made until January, 1968. During the intervening three years the reformers tried to construct a complete and internally consistent "model." Their attention to detail is manifest in the immense corpus of official regulations in which the reforms are embodied, while their concern for coherence led them to introduce the reforms *en bloc* in 1968, having taken only limited transitional measures in 1966-67.

Third, the reforms have throughout had the full political backing of the top level of the Party leadership, for which a price was paid in terms of the alternatives available to the reformers. Although there was extensive public debate on many individual measures and overall enthusiasm certainly varied between members of the Politburo, no broader doubts were expressed openly. The massive propaganda campaign aimed at the middle and lower levels of the Party and the country at large emphasized the unitary development of Party thinking: The reforms of the "economic mechanism" were simply the correct way to adjust the "economic policy" originally established in 1957 to changing circumstances. This guaranteed the stability of Party attitudes toward employment, the standard of living, income dis-

tribution, economic relations with other socialist countries, and desirable directions of structural change in the economy. In turn, however, these given features of "policy" have acted as constraints on possible changes in the "mechanism," while dictating extreme caution with respect to inflation and unemployment, in particular.

Despite these explicit limits, the reforms do make the basic change from a command economy to a socialist market economy. "Approved plans" for "addressees" and the system of plan indices have been eliminated. The plan specifies policy directives only for state administrative bodies. For enterprises it is indicative, and no incentives are attached to the plan in any way. The remaining administrative restrictions on allocation decisions are primarily constraints rather than positive commands.

This shift to a socialist market economy involved no alteration in the form of ownership or management relations, with workers' management in any form having been emphatically rejected. Managers are still responsible to the state administration and the Party. Almost all other determinants of allocation decisions have changed radically. We consider successively the administrative structure, relations between the enterprise and higher authorities, commodity allocation, prices, incentives, working capital, and fixed investment decisions. We then briefly discuss events in the economy in 1968, new measures, and remaining problems, and conclude by evaluating the achievements of the reforms. In addition to demanding some unimportant simplifications, limitations of space preclude any theoretical treatment of decentralization or consideration of the welfare effects of the reforms, the desirable balance between command and market allocation, etc. Our basic concern is the nature and viability of the new system.

The formal partitioning of the economic administration is unchanged, but authority has been redistributed quite substantially, following the shift from commands to other methods of direction. The functional bodies and the enterprises have gained at the expense of the branch ministries, whose diminished role is reflected in personnel cuts of 30-40 percent. Many of these employees followed their jobs to the enterprises.

Among the functional bodies, the major loss

* Since this paper was written, the Hungarian authorities have decreed significant changes in the regulations governing "average-wage control" and the distribution of profit shares, for reasons suggested above. They have also restricted still further the scope of central controls over commodity allocation, but there will be no corresponding relaxation of price controls in 1970.

was sustained by the National Planning Office. Most of its operative authority over investment decisions went to the Investment Bank and other bodies, and its power over material allocation is now exercised by the National Materials and Price Office. (Formerly, when disequilibrium on commodity markets had very little to do with prices, it was just the National Price Office.) The Ministry of Finance, the Ministry of Foreign Trade, and the National Bank are now much more important. Although "control by the forint" has been abandoned, the Bank has discovered that an active short-term credit policy is a more powerful tool.

There has been no "industrial reorganization." Some of the horizontal trusts have been dissolved, but the industrial structure created by the reorganization of 1963 is still highly concentrated (primarily in large enterprises). The few remaining trusts no longer act as a conduit for central commands and are not part of the economic administration. Thus there is no intermediate industrial association of the type seen in other Eastern European countries. Several new "associations" have taken over some service functions from defunct trusts, but they are run by the participating enterprises rather than the reverse and are not business undertakings.

While industrial ministries now have little authority to issue instructions to their enterprises, they still appoint and dismiss the enterprise director and his deputies and determine their base pay. *Ex post* evaluation has replaced operative intervention into enterprise affairs, but there is still some scope for the potentially stultifying influence of ministry control. Recognizing this, the reformers set up for some of the larger enterprises "supervision committees," which are supposed to provide the ministry apparatus with an independent evaluation of enterprise performance. This problem is an aspect of the wider question of ownership and control, and it is too early to say whether the supervision committees will provide any answer to it.

Central physical allocation of commodities has virtually disappeared, although various elements remain in attenuated form. In 1968 there were 35 commodities (covering a few percent of interenterprise turnover) for which at least one user was prescribed a minimum or a maximum amount available to it, but these were only procurement quotas. Users were not required to purchase their full quotas, nor were "output plans" imposed on producers. This contrasts with 1965-66, when about 400 commodities (at the same level of aggregation) were subject to full central allocation.

In 1968 there were also 60 commodities (covering 20 percent of interenterprise turnover) for which monopsonist-monopolist distributors were prescribed.

This limited command authority over allocation is intended to be a temporary means for dealing with particular disequilibria. It is not used to enforce an overall scheme of central priorities or to exert "planners' tension." Central allocation has been replaced by free trade between enterprises. Users (including retail trade) should now have the choice of ordering directly from producers, from wholesalers, or from foreign trade enterprises, although the highly concentrated distributive network will not easily admit active competition.

The foreign trade mechanism has been transformed by changes in both the pricing of traded goods and the contractual relations between foreign trade enterprises and their domestic partners. Foreign prices are converted into domestic prices by separate uniform exchange rates for the dollar and ruble; extensive tariffs and export subsidies in general do not prevent the domestic price from varying directly with the foreign price. Moreover, foreign trade enterprises usually operate on some form of commission. Thus the "price equalization result" and the separation of foreign and domestic markets have been eliminated.

There are still a few aggregate export and import quotas for dollar and ruble markets, enforced by the Ministry of Foreign Trade through its power to deny the requisite export or import permit. But the permit system has apparently proved less bureaucratic than had been feared, its primary use being to regulate the geographical distribution of trade rather than its volume.

Decentralization requires that markets be cleared, at least sufficiently to avoid evoking administrative measures. In particular, the commodity allocation system described above needs a reasonable degree of price flexibility. Of total interenterprise turnover of raw materials and semifinished goods for 1968, about 30 percent was transacted at fixed prices, 40 percent at prices subject to maxima or other limits, and 30 percent at free prices. The corresponding figures for producer prices of finished goods were 3 percent, 19 percent, and 78 percent; for consumer prices, 20 percent, 57 percent, and 23 percent.

The fixed prices and other limits are set by the authorities. As before, in the 1968 producer price reform an average cost pricing schema was used to set these prices; it also provided guidelines to producers for setting the free prices. The average cost figures, based on enterprise cost calculations,

included materials, wages, a 25 percent payroll tax, depreciation, and a 5 percent charge on the gross value of fixed and working capital. To this was added an allowance for profits, computed as a percentage (varying across industries) of the gross value of assets.

Consumer prices are still separated from producer prices by a complex system of turnover taxes and subsidies. Nevertheless, where the consumer price can vary, this affects the producer. Moreover, the number of separate tax and subsidy rates has been cut from about 2,500 outside the garment industry plus 50,000 in it (where the "method of differences" had been used) to about 1,000 over all. Thus producer price ratios now approximate consumer price ratios within much broader product groups than before.

The incentive system for the enterprise is based on its profits. Despite the complexity of the system, it does introduce true profit taxation—the budget no longer takes the residual after certain allowances to the enterprise. From gross revenues are deducted costs, which here include materials, wages, depreciation, payroll tax, capital charge (in 1968 paid only on assets owned by the enterprise; i.e., not financed by credit), and interest on short- and long-term credit. The enterprise may then pay a further tax on gross revenues (or some part thereof) or receive subsidies. The resulting amount is taxable profit (subject to an important modification discussed below). This is divided into development and sharing portions in proportion to the capital-labor ratio of the enterprise. The former is taxed at 60 percent, the latter progressively (the degree of progressivity being rather steep), leaving the development and sharing funds. A small part of each goes into a reserve fund. The enterprise may use the remainder of the development fund (plus 60 percent of depreciation allowances, on average) to expand fixed and working capital, while the sharing fund finances welfare expenditures and cash distributions to workers and staff. The rules governing these distributions make them a very substantial part of the incomes of executives (including the director and his deputies) and higher-level technical employees, but a much smaller part of manual workers' incomes.

A key element of this system is "average wage control." On the one hand, before taxable profit is divided into sharing and development portions, it is augmented (for accounting purposes only) by that part of the total wage bill attributable to the excess of the average wage during the year over the "base" average wage level of 1967 (i.e., employment times the difference between current

and base average wage). On the other hand, this same sum is deducted from the post-tax sharing fund. In effect, all wage increases after 1967 are paid out of the post-tax sharing fund before it is distributed—and so is any increase in the average wage which might result from a change in the composition of the enterprise's labor force. Average wage control is intended to combat cost inflation and provide a microeconomic force against unemployment. It is clearly a strong deterrent to wage increases and an incentive to retain, even to seek, low-wage labor. An increase in the enterprise average wage will normally reduce total managerial incomes, because these are so sensitive to the amount in the sharing fund available for distribution.

The taxes and subsidies affecting the total of taxable profits are differentiated as between enterprises. Some enterprises pay a "production tax," at rates and on bases which differ between enterprises. This tax is intended to draw off rents, "excessive" quasi-rents (from having much better equipment than other enterprises), or "excessive" export profits (although a large proportion of exports may make losses, the range of efficiency is so wide that others may make very high profits). Of the subsidies, a few are direct payments to enterprises which would otherwise run at a deficit (the amount of the subsidy being fixed for several years at a time). The remainder are export subsidies, usually given as a percentage of export revenues. These are necessary because of the considerable overvaluation of the forint in relation to both the dollar and the ruble; they apply to well over half the total value of exports.

The division of taxable profits according to the enterprise capital-labor ratio (where labor is generally represented by twice the wage bill) is an attempt to find a fair way of determining automatically the enterprise's retention ratio. The need to do this is another aspect of the ownership-control problem—neither managers, nor workers, nor the ministries could be allowed (for different reasons) to fix the retention ratio in each enterprise individually.

With only minor exceptions, the rules of the profit taxation system are uniform between enterprises (given the subsidies and production tax). They were set in the first instance for three years, in order to preclude annual "equalizing" (the "ratchet effect") and allow enterprises to plan ahead with confidence that these rules will remain unchanged. Yet in attempting to encourage longer horizons in this way, the authorities have denied themselves important instruments of fiscal policy.

The reforms introduced true self-financing.

After a "working capital settlement" on January 1, 1968, on average about 75 percent of the working capital of industrial enterprises was "own funds," the remaining 25 percent being financed by credit. Any permanent increase in working capital was to come out of the development funds or medium-term credit repayable only from development funds, while temporary working capital needs could be financed by short-run (less than one year) credits repayable from gross revenues. The 1968 interest rate on both types of credit was set at 8 percent. Trade credit, while still strongly disapproved by the authorities, was made technically possible by giving the enterprise free control over its own bank account. The Bank no longer effects payment automatically at the initiative of the seller.

There are now three major sources of finance for fixed investment: budgetary funds, credits issued by the Investment Bank (repayable from development funds), and development funds. But the principle that almost every productive investment, even if initiated by the planners, should be at least partially financed by the enterprise entails that the patterns of finance and decision making are no longer congruent. Although the enterprises should eventually obtain significant investment decision authority, the large carryover of unfinished projects from the prereform period has delayed this transfer. The enterprises must use a good part of their development funds and credit to finance these projects. Enterprises do already have more scope for independent investment decisions, however, and they now have incentives to minimize costs and completion times.

The effect of credit policy on investment is not yet very important, simply because there is relatively little room for credit, given the size of the development funds and of state investment demand. A new policy instrument is the import deposit required on imports of machinery from convertible currency markets; in 1968, 150 percent of the purchase price had to be deposited (from the development fund) in a noninterest-bearing account at the Bank, repayable after two years.

The capital market is still limited in many respects. For example, there is no analogue of true long-term credit—debentures issued to finance significant expansion. (In 1968, long-term credit issued by the Investment Bank had an average maturity of four years.) This kind of investment is reserved to the state: in effect, it centralizes the necessary funds and lends them out at 5 percent; the capital charge is a fixed prior

charge, like interest on debentures. Moreover, there is no explicit provision for interenterprise capital flows. Enterprises can in principle pool assets to set up a "common enterprise" in which each of the founding enterprises holds shares, but there are few instances of this so far. It is certainly unlikely that the shares could be bought and sold.

With this brief description of the reforms, we can now survey events in the Hungarian economy in 1968 and the reformers' subsequent response. Industrial output rose 5 percent, only slightly less than planned, but industrial employment rose 4 percent, so that output per man rose much less than expected. Some favorable but limited changes in the structure of output and its adjustment to demand were noted, mainly in more rapidly growing branches of heavy industry (telecommunications equipment, precision instruments, chemicals). Several of the most significant changes, however, were consequences of centrally-determined patterns of development fixed in the Five-Year Plan (aluminum, transportation equipment).

The increase in stocks of goods of industrial origin was even greater than in 1967, when inventories had consciously been built up at a record rate in order to forestall any serious disequilibria in the first year of the reforms. Producers responded to the relaxation of allocation controls by expanding input stocks (especially of metallurgical goods) very rapidly, while wholesale trade accumulated stocks of some consumer goods (especially clothing) because of deficient demand.

Both producing and trade enterprises had little difficulty in financing this stockpiling. Short-term credit policy was permissive, in order to ease the problems of transition to the new system. Enterprises were not deterred by interest costs, and they were in any case very liquid, because profits were over 20 percent greater than expected. This excess should not have been surprising: it was primarily the result of setting prices, taxes and subsidies on the basis of enterprise cost estimates, which of course contained ample safety margins.

Perhaps because these margins were so high, industrial producer prices seem to have been remarkably stable during the year, rising perhaps 1–2 percent (one does not know how much to trust the statistics here). Excess demand in some markets often did not pull prices up, in part because disequilibria were most frequent in fixed-price markets. Construction prices rose 4–5 percent more than planned, however, mainly in response to excess demand. Investment demand remained strong, with total investment up 2 per-

cent in real terms (as planned) over 1967, a boom year. Investment demand shifted toward construction; the relative increase in machinery prices in the price reform was quite large (in particular for imported machinery, especially machinery purchased with convertible currencies).

Total personal money incomes went up 8 percent, mostly because of the rise in employment. Average wages paid in industry rose only 2 percent, but profit shares (paid in March, 1969) rose much more, so that total per capita payments to industrial employees rose 3-4 percent. Personal savings rose considerably, taking some of the pressure off the consumer goods market. Thus the consumer price index rose only 2 percent during the year.

The balance of trade improved somewhat over all, mainly because of an increase of 10 percent in exports to socialist countries. Imports from these countries rose only 4 percent, and a substantial surplus replaced the small deficit of 1967. The entire surplus was attributable to trade with Czechoslovakia and the U.S.S.R.; the share of the latter in Hungarian foreign trade rose to 38 percent of exports, 36 percent of imports. The balance with nonsocialist countries deteriorated slightly, since a small improvement with developed Western countries (exports to which fell 2 percent, imports 5 percent) was outweighed by a switch into deficit with developing countries.

The new measures taken for 1969 in response to these developments are revealing of the authorities' attitudes. They are willing to trust the system and to await further evidence, even in the face of some disturbing phenomena. Rather than trying to correct deviations from the plan, they are attempting to adjust to them. For the most part they have gone ahead with their announced intentions to loosen controls still further.

The proportions of total interenterprise turnover covered by quotas and statutory monopolist-monopolist distributors have each been cut roughly by half. More turnover tax rates have been unified, and the proportion of retail turnover covered by free consumer prices has been raised to 30 percent. Producer prices have been made somewhat more flexible as well. The import deposit on machinery from convertible currency markets has been lowered to 100 percent. Enterprises had been told before 1968 how their subsidies would be cut from year to year, and the total amount of subsidies paid has in fact been reduced by 10 percent. Many tariffs had been suspended in 1968 because enterprises pleaded inability to bear the increases in input prices; these are now in force.

Certain policy instruments have been used to deal with particular problems revealed by the economy's performance in 1968. In an attempt to reduce stock accumulation, quantitative credit controls have been tightened and effective interest rates have been raised by eliminating the exemption of credit-financed assets from the capital charge. Here the authorities may have overreacted; this amounted to a real credit squeeze, as they have discovered. New tax incentives have been introduced for the expansion of relatively profitable exports to convertible currency markets. They were at least partly responsible for the increase of 32 percent in convertible currency exports during the first nine months of 1969 (over the corresponding period in 1968). A 6 percent tax on construction activity is intended to cut demand and skim off some excess profits in this sector.

What has not been done is equally significant. For example, decentralized investment rose more than planned in 1968, and new accruals to development funds (for 1969) were inflated by excess 1968 profits. But there have been no special levies, and the authorities were willing to make room for more enterprise investment by cutting back on intended new starts of large state projects. New regulations will have the effect of increasing enterprise contributions to large projects, but it appears that they will not reduce the proportion of decentralized investment.

There are several unresolved problems. First and most urgent is the apparent sharp slowdown in the growth of industrial output so far this year (again, the statistics are suspect). This has been accompanied by a large increase in industrial employment. The record harvest does allow the planners to wait a bit before acting, and forecasts indicate some improvement may be expected. But producers in many branches of industry seem to be facing very weak demand—not final demand, since exports, consumption and investment are holding up well, but demand from the trade network and possibly from other producers. Liquidity fell drastically because of the credit squeeze, the loss of retained profits consequent upon distribution of the sharing funds, and a new requirement that reserve funds be deposited in special accounts rather than used to finance working capital. The apparent result has been a substantial absolute fall in stockholding. Wholesale trade had already started to cut orders in late 1968. Those enterprises which could find export markets did so; others simply had to reduce output.

One might in any case question the wisdom of restricting credit to trade enterprises, since the

drop in their stocks will hinder the development of the distributive network and may resurrect the shortage mentality of the past in producers. Some increase in stockholding was to be expected and welcomed after allocation controls were relaxed, although it should have been accompanied by price reductions to dispose of otherwise unsalable stocks.

Perhaps the most worrisome aspect of these events is the concomitant rise in employment. One tends to attribute this to average wage control, but there are complications. First, the working week has been reduced from forty-eight to forty-four hours for successive groups of enterprises, beginning in 1968. But the data for 1968 show that these enterprises had much better output and output per man performance than those not yet affected, and they did this without raising overtime. Another possible explanation is that the entire increase in industrial employment has taken place outside Budapest, because of the relocation of enterprises to the provinces and expansion of those already there; and new laborers from rural areas will be less productive initially than urban workers. There may also have been a shift toward more labor intensive products. Whatever the role of average-wage control in the disappointing productivity figures, the planners are understandably reluctant to give up the constraint it imposes on wage inflation, especially since many workers already feel the impact of consumer price increases. The planners may also fear provoking unemployment. The balance of the argument is not yet clear.

Another pressing issue is public discontent with the rules for distribution of profit shares. The large sums going in a single, visible lump to managers do not commend the system to workers, who are beginning to realize why managers have been so reluctant to increase wages. Since profit shares have replaced managerial premia, total managerial incomes actually increased by only 4 percent in 1968 (they fell by 7 percent in metallurgy and engineering), but this is not so easily seen by the workers. Thus the system has created ill-feeling without even accomplishing one of its prime objectives: an increase in earnings differentials. Moreover, tying managerial incomes so closely to annual profits may in itself be unwise, insofar as managers may become preoccupied with short-run profits to the neglect of long-run considerations. Finally, one may doubt the rationale of basing incentives so heavily on profits when the price system is still such an imperfect construction.

The price system itself will require reconsidera-

tion in the near future. It is hard to tell from available data how prices could seemingly have remained so sticky—both downward and upward—without more widespread evidence of disequilibria. The fall in demand discussed above has revealed some elasticity in producers' supply curves but no willingness of producers to test the elasticities of the demand curves facing them. Enterprises are of course unaccustomed to operating independent price policies. The aggregative indices, into which a significant proportion of fixed prices enter, may simply mask widespread price flexibility among free-price goods, or contrary movements in these prices may cancel each other out. But the data show virtually no change in interindustry price relatives since January, 1968.

Whatever is done about price flexibility, a large proportion of prices will remain controlled for some time, while existing interindustry and intersectoral price ratios give little guide to efficient allocation. The substantial variation of profit rates between industries has little apparent justification, and the massive subsidies cloud the picture still further. Subsidies (primarily export subsidies) received by industrial enterprises in 1968 represented 50 percent of total profits in industry. These seem to have been concentrated on a small group of big enterprises. Other very large items are the subsidies to consumer prices, to the prices of inputs into agriculture (necessary because agricultural output prices are still artificially low), and direct subsidies to agricultural cooperatives. It is intended to reduce all of these over time, but the pace may be very slow. In the interim, costs and prices are heavily distorted throughout the economy.

The inadequacies of the price system are an important obstacle to the establishment of freer capital movement, flexible forms of cooperation between enterprises, and some kind of capital market. Capital cannot be allowed to flow freely in response to profit rate differentials if these give misleading information. The absence of a capital market is felt all the more because the only source of credit is the monopolized banking system. There is little doubt that this monopoly must eventually be broken. As it stands, the Bank is hardly an appropriate vehicle to encourage risk-taking and imaginative ventures.

The high degree of concentration in industry will not change in the foreseeable future. There has been some emphasis on the virtues of decentralization within existing large firms, but so far little has been done in this direction.

Finally, we should mention Hungary's desire

to change the mechanism of its trade relations with its Comecon partners. The reforms in foreign trade do not fit well with current Comecon trading arrangements; the large, highly differentiated export subsidies were in part motivated by the need to induce enterprises to fulfill quotas fixed in Comecon negotiations without direct instructions. This is not a long-term solution, but the conservative nature of the reforms in other countries has meant that they do not feel the same urgency about these problems as does Hungary.

The two years under the reforms have proven that the economy can operate without commands. Enterprises have been able to decide from whom to purchase inputs, what to produce, and to whom to sell, all without chaos—indeed, with a reasonable degree of equilibrium. The transition has been accomplished without inflation, without disastrous effects on the balance of payments, and without unemployment. The legacies (using Professor Neuberger's apt term) of the command system are still powerful. In particular, enterprise horizons are still very short: they are unwilling to take risks, they do not innovate at a satisfactory rate, and they have not fully explored the range of their new authority. Moreover, the immense distance of the previous price-input-output pattern from any kind of equilibrium has required, in order to facilitate the transition, various measures which have undesired effects: a wide range of direct controls, some degree of differentiation and specificity (to individual enterprises) of the newly operating parametric controls, and innumerable distortions in the price system (reflecting distributional and other constraints). But the caution which dictated these measures was sensible, and

all are amenable to correction within the basic context of the system.

Our major conclusion is that there is little prospect of recentralization. The planners have shown themselves willing to accept enterprise actions differing from the plan. Moreover, while imposing limitations on their own discretionary authority, they have so restructured the apparatus of economic administration that there is very little opportunity for creeping recentralization.

Recentralization would therefore require a conscious, open political decision. At present there seems little likelihood of pressure for such a decision arising within the country. To be sure, the trade unions are criticizing various aspects of the reforms and their effects on workers, and there are conflicts between managers and Party secretaries. But the public commitment of the entire top level of the Party to making the new system work is very strong. The obvious possible source of external pressure is the Soviet Union. Hungary has taken great care to make the reforms palatable to the Soviet Union. It has more than kept up its end of Comecon trade, while increasing its dependence on trade with the U.S.S.R. Its overtures to the West have been cautious; although it has obtained convertible currency loans, they have not been excessively large or widely publicized. It has made its aversion to workers' management unmistakably clear. And for its part, the Soviet Union has taken an approving attitude—not wildly enthusiastic, but certainly not critical, and colored by some genuine curiosity as to how it will all work out. This seems a reasonable position to maintain.

GROWTH PERFORMANCE OF EASTERN EUROPEAN ECONOMIES AND COMPARABLE WESTERN EUROPEAN COUNTRIES*

By BELA BALASSA and TRENT J. BERTRAND
Johns Hopkins University

I

The purpose of this paper is to compare the growth performance of the centrally planned economies of Eastern Europe with that of European economies characterized by decentralized decision making. The investigation covers five centrally planned economies (Bulgaria, Czechoslovakia, Hungary, Poland, and Rumania), four Western European countries (Spain, Greece, Norway, and Ireland), and Yugoslavia which has developed a decentralized socialist system. The choice of the countries included in the study reflects a desire to have a representative group operating under different economic systems at comparable levels of development although it has, to some extent, been constrained by the availability of the data.

The study covers the time period 1953 to 1965; the initial year has been selected to avoid the effects of postwar reconstruction in several of the countries in question while the choice of the terminal year has been dictated by data availabilities. The investigation is limited to manufacturing industry where the most important structural changes have taken place.

As an indication of the relative level of economic development of the countries in question, in Table 1 we present estimates of per capita gross national product calculated by the use of purchasing power parities for the year 1955. The results may be subject to considerable error and, due to differences in the statistical sources, there is greater comparability in the data among the centralized and among the decentralized economies than between countries that belong to different groups. Nonetheless, it is encouraging that in the case of Yugoslavia—the only country that appears in both sets of calculations—the two estimates are very close. In any case, it would appear that the countries under study can be classified into three groups according to their relative levels of economic development: low—Greece, Yugoslavia, Bulgaria, Spain and Rumania; intermediate—Ireland, Poland, and Hungary; and high—Czechoslovakia and Norway. The latter two groups will further be combined in

* The authors would like to acknowledge assistance from Michael Claudon with the research for this paper.

TABLE 1

ESTIMATES OF GROSS NATIONAL PRODUCT PER CAPITA
OF SELECTED EUROPEAN COUNTRIES IN U.S. DOLLARS,
1955, DERIVED BY THE USE OF PURCHASING
POWER PARITIES

	(1)	(2)
Bulgaria.....	515	
Czechoslovakia.....	1,163	
Hungary.....	835	
Poland.....	755	
Rumania.....	551	
Greece.....		336
Ireland.....		704
Norway.....		1,394
Spain.....		516
Yugoslavia.....	444	427

SOURCE: Column 1 derived from data in F. L. Pryor and G. J. Staller, "The Dollar Values of the Gross National Products in Eastern Europe, 1955," *Economics of Planning*, Vol. 6, No. 1 (1966). Column 2 from Bela Balassa, *Trade Prospects for Developing Economies* (Richard D. Irwin, Inc., 1964), pp. 384 and 385.

distinguishing between "less developed" and "more developed" countries among those under study.

II

International comparison of growth performance is a formidable task and we approach it with some trepidation. There are difficulties in obtaining comparable data for the different countries, and a summary measure cannot take account of all the relevant variables. Comparisons of growth rates of output disregard changes in inputs, marginal capital-output ratios neglect labor as well as capital accumulation in earlier periods, while labor productivity comparisons abstract from the capital factor.

In recent years, it has become popular to separate the growth of output attributable to (a) the growth of factor inputs and (b) a residual presumably reflecting organizational changes, improvements in labor quality, and technical progress. The residual—also called the growth of

total factor productivity or, simply, total factor productivity—has been taken by some writers to be an indicator of growth performance [1] [2].

The separation of factor inputs and total factor productivity as sources of economic growth assumes that (a) there are no economies of scale; (b) organizational changes, improvements in labor quality, and technical progress are independent of the growth of factor inputs; and (c) factor inputs represent purely a cost and their growth does not contribute to economic welfare. We will initially accept these assumptions in calculating total factor productivity for the countries under study. Subsequently, we will remove assumptions (a) and (b) in order to estimate net factor productivity, defined as the residual growth rate that is obtained if one adjusts for both the direct and indirect effects on growth of increases in factor inputs. Finally, we will consider the question of the welfare significance of increases in factor supplies.

III

In calculating total factor productivity, we have first attempted to fit production functions of the form shown in equation (1) to data on indus-

$$(1) \quad O = \sigma^{\gamma} L^{\alpha} K^{\beta}$$

trial output (O), employment (L), and the capital stock (K) in manufacturing industries, both with and without constraining the sum of the exponents to one. However, for various reasons, such as multicollinearity, the shortness of the period of observation, and the variability of the data, statistically significant results have not been obtained. We have therefore used equation (2), obtained by differentiating the logarithmic form of equation (1), to estimate total factor productivity (γ).

In the equation, σ , l , and k denote the rate of growth of industrial output, labor, and capital

$$(2) \quad O = \gamma + \alpha l + \beta k$$

respectively, while α and β are the elasticities of output with respect to labor and capital. With perfect factor markets, α and β are the shares of labor and capital, and we can use estimates of factor shares to weight the factor inputs. But there are no capital markets in the Eastern European economies, and the valuation of capital is largely arbitrary. Accordingly, it has been necessary to make some assumptions concerning the rate of return to capital in these countries.

In the Western European countries under study, the share of labor is between 45 and 65 percent and the gross rate of return to capital ranges from 18 to 22 percent. If 5 percent is

deducted for depreciation, the corresponding net rate of return is between 13 and 17 percent.¹ These rates have been used in evaluating the returns to capital in the Eastern European countries for which capital stock data are available in absolute terms: Hungary, Rumania, Czechoslovakia, and Yugoslavia. The resulting labor shares for these countries are all in the 45 to 65 percent range. Given the uncertainties associated with the data, we have used both 45 and 65 percent as an estimate of the share of labor in manufacturing industry.

Data on the growth of output, capital, and labor are presented in columns 1 to 3 of Table 2 while estimates of total factor productivity calculated under alternative assumptions as regards labor's share in output are shown in columns 4 and 5. To avoid giving undue weight to initial or terminal years, growth rates of output and factor inputs have been calculated by regressing the relevant data on time. Growth rates of output and capital stock are in real terms although there is a suspicion that the Bulgarian capital stock figures have not been appropriately deflated. Data on the capital stock and, to a lesser extent, on output are subject to considerable error in the other socialist countries, too, and error possibilities exist also in regard to capital stock estimates for Western European countries. Sources of data are described in the Appendix.

The interpretation of the results is reasonably straightforward. Irrespective of the choice of input weights, Rumania, Spain, and Yugoslavia show the most rapid increases in total factor productivity among the countries in question. With an appropriate deflation of the capital stock figures, Bulgaria would probably also appear in this group. The results thus lend credence to the contention that among countries that have passed the take-off point, the possibilities for utilizing existing technological knowledge permit more rapid growth in countries at lower than at higher levels of development.

Greece, however, provides an exception inasmuch as it belongs to the first group of countries in terms of the level of development and the second in terms of growth performance. Within the latter group, Norway and Czechoslovakia appear to be in the lead, followed by Greece and Poland, with Ireland and Hungary at the bottom of the list. But the ranking of these countries is affected if different input labor shares are assumed in particular cases.

¹ By comparison, a range of 8 to 20 percent has been suggested by Bergson in regard to the Soviet Union [2, p. 20].

TABLE 2

RATES OF GROWTH OF OUTPUT AND FACTOR INPUTS AND MEASURES OF FACTOR PRODUCTIVITY
IN THE MANUFACTURING INDUSTRY OF SELECTED EUROPEAN COUNTRIES

Country	Period	(1) (2) (3) Annual Rate of Growth			(4) (5) Total Factor		(6) (7) (8) (9) Net Factor Productivity			
		Output	Capital	Labor	Productivity		Case A		Case B	
					$\alpha = .45$	$\alpha = .65$	$\alpha = .45$	$\alpha = .65$	$\alpha = .45$	$\alpha = .65$
Bulgaria.....	1953-65	12.5	11.6	7.6	2.6	3.4	— .3	.7	—2.3	—1.5
Czechoslovakia.....	1953-65	7.0	6.6	2.7	2.1	2.9	.7	1.6	— .3	.6
Hungary.....	1953-65	6.5	7.3	3.0	1.1	2.0	— .4	.6	—1.5	— .2
Poland.....	1961-65	6.6	6.5	3.0	1.7	2.4	.2	1.1	— .8	.2
Rumania.....	1953-65	11.1	8.3	4.1	4.6	5.5	2.7	3.8	1.4	2.6
Centrally planned economies										
—including Bulgaria*...					2.4	3.0	.6	1.5	— .7	.3
—excluding Bulgaria*...					2.4	3.2	.8	1.7	— .3	.7
Greece.....	1951-65	6.9	7.1	2.8	1.7	2.6	.2	1.3	— .8	.4
Ireland.....	1953-65	4.7	4.2	1.7	1.7	2.2	.8	1.4	.1	.9
Norway.....	1953-65	5.4	5.1	0.8	2.2	3.1	.7	2.4	.6	1.9
Spain.....	1959-65	11.2	8.7	4.5	4.5	5.3	2.5	3.7	.3	1.6
Decentralized private enterprise economies*					2.5	3.3	1.0	2.2	.1	1.2
Yugoslavia.....	1953-65	11.8	7.5	6.7	4.5	4.7	2.4	2.6	1.0	1.2
Decentralized economies*					2.9	3.6	1.3	2.3	.3	1.2

* Unweighted average.

SOURCES: See Appendix.

It further appears that unweighted averages of total factor productivity differ little between the centralized socialist economies and the decentralized private enterprise economies. The situation changes, however, if we add Yugoslavia to the second group. Now decentralized economies seem to be ahead of centralized economies by at least one-half of one percentage point.

IV

Using total factor productivity as an indicator of the country's growth performance reflects the assumption that the results are independent of the rate of growth of the inputs themselves. This assumption will not be fulfilled if the quality of productive factors of a later "vintage" is superior to those of an earlier vintage or if there are increasing returns to scale. In the first eventuality, the average quality of factors will be positively related to the rate of growth of factor inputs; in the second, the rapid growth of factors will bring

forth increases in output by increasing the scale of production. As to the former, the reduction in illiteracy, improvements in schooling, and the institution of training programs which may be associated with the rapid growth of labor in manufacturing, will raise the quality of the labor force while the capital stock is upgraded as new investments embody more advanced technology.

We have attempted, therefore, to divide the estimates of total factor productivity into two parts: a part explained by the indirect effects of the growth of productive factors and a residual termed "net factor productivity" which may more adequately reflect the differences between countries in regard to their systems of economic decision making and organization. To do this, it is necessary to quantify the relation between factor input growth and total factor productivity as defined earlier.

For U.S. manufacturing, A. A. Walters has found evidence of increasing returns to scale of

from 27 to 35 percent [5]. But Walters' results may reflect improvements in factor qualities as well, and we have interpreted them as such. At the same time, one can assume that in countries at lower levels of development there is more scope for improving factor qualities and exploiting economies of scale than in the United States. The results obtained for U.S. manufacturing industries therefore may represent a lower limit to the joint effects of the two influences as they operate in national economies that are less developed industrially.

At any rate, the U.S. results derive from the experience of a single country, and greater confidence can be placed in the figures if inter-country comparisons are made. We have attempted to do this by estimating the relationship between the average growth of factor inputs and total factor productivity in a cross-country regression. In order to increase the number of observations available, we have used data for four-year subperiods in the countries under study, with the exclusion of Bulgaria whose capital stock figures are suspect. The results are given in equations (3) and (4).

$$(3) \quad \gamma_1 = .09 + .54T_1 \quad R^2 = .50, \quad D.W. = 2.27 \\ (0.13) (4.24)$$

$$(4) \quad \gamma_2 = 1.60 + .45T_2 \quad R^2 = .47, \quad D.W. = 1.61 \\ (2.76) (3.96)$$

where T_1 and T_2 are the growth rates of factor inputs weighted by $\alpha = .45, \beta = .55$ and by $\alpha = .65, \beta = .35$, respectively, and t values are shown in parentheses.

The results tend to confirm our suspicion that in countries at a level of development as those under study, the indirect effects of growth in factor inputs on the growth of output can be very substantial. Equations (3) and (4) show this to be from 45 to 54 percent of the direct effect. It should be added that the t values correspond to levels of significance exceeding 99 percent while the exclusion of other variables from the equation limits the values taken by the coefficients of determination. On the basis of Walters' results and our findings, we have next adjusted the total factor productivity figures on the assumption that a 1 percent increase in combined factor inputs yields (A) an added .3 percent or (B) an added .5 percentage point growth of industrial output. The resulting net factor productivity figures are presented in columns 6-7 and 8-9 of Table 2.

This adjustment favors countries where the rate of growth of factor inputs has been relatively

low. Thus, Norway joins Rumania, Spain, and Yugoslavia among countries with the highest net factor productivity and the relative positions of Ireland and Czechoslovakia are also improved. Furthermore, the imbalance between the less developed and the more developed countries of the group indicated by comparisons of total factor productivity is greatly reduced. For instance, while the unweighted average of total factor productivity (with $\alpha = .45$) is 5.0 for Greece, Spain, Yugoslavia, and Rumania as compared to 3.0 for Czechoslovakia and Norway, the corresponding net factor productivities are 2.8 and 2.0 under case A and 1.4 and 1.2 under case B.

Using net factor productivity as an indicator of growth performance, the differences between Yugoslavia and the other decentralized countries also tend to diminish—significantly so if the higher labor share is assumed. In turn, decentralized economies now appear in a much more favorable light than the centrally planned economies. While the results for the latter group of countries are affected by the exceptionally poor showing of Bulgaria due in part to the previously noted data problems, even excluding Bulgaria net factor productivity is decidedly higher in decentralized than in centrally planned economies. Thus, the hypothesis that the comparable gains in total factor productivity attained in Eastern Europe are the result of more rapid growth in factor inputs appears to be supported by our findings.

Finally, it should be noted that factors other than the direct and indirect effects of input growth would seem to have a negative impact on the growth of output in five countries if the lower (.45) value is assumed for the α coefficient and the higher value (.5 percent) is used for indirect effects. This unlikely result may be interpreted as evidence that the higher values of α are more reasonable. With $\alpha = .65$ only Hungary and Bulgaria show negative net factor productivity, and Hungary is a special case because of the 1956 events, as is Bulgaria because of data problems. Of course, the results might also be interpreted as evidence that the indirect effects of factor input growth on total factor productivity are less pronounced than is assumed in case B.

V

Net factor productivity has been estimated under the assumption that part of the increase in total factor productivity is explainable by the indirect effects of the growth of factor inputs on output growth, when such indirect effects reflect economies of scale and improvements in factor

qualities. There are few problems with the first assumption, but the second opens possibilities of error since it disregards possible differences in the rate of improvement of the quality of capital and labor. The resulting error will be the greater, the larger are differences in factor growth rates.

To remove this source of error, it would be necessary to estimate the extent of improvements in the capital stock and in labor over time. While "embodied" technical progress has been much discussed in recent years,³ no reliable estimates exist and even less can be said of its numerical importance in the countries under consideration. Similar considerations apply to the extent of improvements in the labor force. At the same time, Walters' findings and our regression results suggest that the proportionality assumption made in this study does not involve substantial error.

The problem remains that we have considered factor inputs purely as a cost and thus have attached no welfare significance to increasing the rate of saving and reallocating labor from low-productivity occupations or from the ranks of the unemployed to manufacturing industry. Yet, while in Western Europe (although not in centrally planned economies) capital accumulation largely depends on individual decisions and the absorption of labor in manufacturing industry is predicated on the availability of labor, the growth of factor inputs is affected by the economic policies followed.

It has not been possible to incorporate such considerations in the present study, however. Apart from the difficulties of statistical measurement, these considerations raise intricate welfare problems. In this connection it may be recalled that the Soviet Union was strongly criticized for forced reductions in consumption (increasing the rate of saving) and forced collectivization (increasing the mobility of labor) during the 1930's. Today some would look differently at that period of "initial accumulation" but any judgment would involve interpersonal and intertemporal comparisons which we have wished to avoid in this paper.

VI

A further question is whether there have been significant changes over time in the growth performance of the countries under consideration. In order to deal with this question, we have divided the total period into two subperiods—1953–59 and 1959–65. But instead of simply comparing estimates of factor productivity in the two subperiods, we have used a statistical test which

³ This problem was first raised in [4].

involved calculating the value of the elasticity of output with respect to labor (α) that would equate factor productivity between periods in the three cases considered earlier.³ A judgment as to whether the calculated values of α are reasonable or not permits us to conclude whether and in what direction factor productivity has changed.

In Table 3 we show the condition placed on α in order that factor productivity in the period 1959–65 is greater or equal to that for the period 1953–59. On the assumption that the actual value of α will be between 0 and 1 and, most likely, between .45 and .65, we can conclude that (a) an increase in factor productivity has definitely occurred in Spain, Hungary, and Ireland while there has been a decline in Czechoslovakia and Greece and (b) factor productivity also seems to have increased in Bulgaria and Norway and decreased in Rumania and Yugoslavia.

There is some support in our findings for the hypothesis that in the more developed economies of the group, decentralized decision making has been more successful in maintaining or raising factor productivity than central planning. Thus, an increase in factor productivity has occurred in Norway and Ireland while a decrease is shown for Czechoslovakia. The exception to this hypothesis provided by Hungary may in fact be just that: an exception explainable in terms of the retarding effects in the earlier period of the 1956 events. As far as the less developed countries are concerned, no unambiguous conclusions are suggested by the results since improved performance occurs in Spain and Bulgaria, while a decrease in total factor productivity is shown in Greece, Rumania, and Yugoslavia.

VII

In this paper an attempt has been made to compare the growth performance of five centrally planned economies, four Western European countries and Yugoslavia—the only country with a decentralized socialist system. The countries selected for the investigation can also be classified according to the degree of economic development: low (Greece, Yugoslavia, Bulgaria, Spain, and Rumania), intermediate (Ireland, Poland, and Hungary), and high (Czechoslovakia and Norway).

Calculations of total factor productivity for the manufacturing sector in the 1953–65 period show the low-income countries other than Greece in the lead, with Ireland and Hungary at the bottom of

³ This technique of analysis is due to Kaplan and has been previously used by him to analyze changes in total factor productivity in the Soviet Union [3].

TABLE 3

CONDITIONAL VALUES OF THE ELASTICITY OF OUTPUT WITH RESPECT TO LABOR FOR WHICH TOTAL FACTOR PRODUCTIVITY INDUSTRY IN 1959-65 IS AS GREAT OR GREATER THAN IN 1953-59 FOR SELECTED EUROPEAN COUNTRIES

Country	Comparison of Total Factor Productivity	Comparison of Net Factor Productivity	
		Case A	Case B
Bulgaria.....	$\alpha \geq .34$	$\alpha \geq .31$	$\alpha \geq .30$
Czechoslovakia.....	$\alpha \geq 4.03$	$\alpha \geq 3.23$	$\alpha \geq 2.88$
Hungary.....	$\alpha \leq 1.86$	$\alpha \leq 1.20$	$\alpha \leq 1.01$
Poland.....	(n.a.)	(n.a.)	(n.a.)
Rumania.....	$\alpha \leq .36$	$\alpha \leq .10$	$\alpha \leq .01$
Greece*.....	$\alpha \leq .96$	$\alpha \leq .46$	$\alpha \leq .23$
Ireland.....	$\alpha \leq 12.47$	$\alpha \leq 4.65$	$\alpha \leq 1.66$
Norway.....	$\alpha \leq 1.15$	$\alpha \leq .92$	$\alpha \leq .78$
Spain†.....	$\alpha \leq 1.86$	$\alpha \leq 1.34$	$\alpha \leq 1.11$
Yugoslavia.....	$\alpha \geq .98$	$\alpha \geq .82$	$\alpha \geq .75$

* 1961-65 compared with 1951-61.

† 1961-65 compared with 1957-61.

the list. Averages of total factor productivity for centrally planned economies and Western European countries are about the same but adding Yugoslavia to the second group puts decentralized economies ahead of centrally planned economies.

The measurement of total factor productivity reflects the assumptions of constant returns to scale and unchanged factor qualities over time. These assumptions are not fulfilled in practice; indeed, our results indicate the existence of a positive correlation between the growth of factor inputs and total factor productivity. Adjusting for the effects of the former on the latter, we obtain estimates of net factor productivity. As this adjustment favors countries with a relatively low rate of growth of factor inputs in manufacturing, the relative positions of Norway, Ireland, and Czechoslovakia improve while Bulgaria, Spain, and Yugoslavia now rank lower.

It also appears that if we adjust for the indirect effects of the growth of factor inputs, differences in the growth performance of countries at lower and at higher levels of development are reduced; Yugoslavia's advantage over decentralized free enterprise economies disappears; and the latter group of countries has a decided edge over centrally planned economies. Such an advantage is also shown if we consider changes in factor productivity over time.

Apart from the conceptual problems and the statistical difficulties of the calculations, note

should be taken of the fact that the procedure applied considers the growth of factor inputs as a cost and attaches no welfare significance to it. If instead we take a high rate of saving and increases in manufacturing employment to have welfare significance in themselves, the relative position of the socialist countries will improve. But against this we should set the relatively poor performance of agriculture in these countries which has not been considered in this study; yet increases in the capital stock and in labor have taken place in part at the expense of agriculture.

APPENDIX

Bulgaria: Growth rates of industrial output, capital stock and labor were derived from Lazarcik and Wrynyczuk, *Bulgaria: Growth of Industrial Output 1934 and 1948-1965*, Occasional Paper No. 27 of the Research Project on National Income in East Central Europe, pp. 4 and 9.

Czechoslovakia: Growth rates of industrial output, capital stock, and labor were derived from Lazarcik and Staller, *A New Index of Czechoslovak Industrial Output 1937 and 1947-1965*, Occasional Paper No. 24 of the Research Project on National Income in East Central Europe, pp. 24 and 27. In estimating factor shares, payments to labor were obtained from Lazarcik and Staller, *ibid.*, p. 49, while value of net capital stock was estimated as 75 percent of full replacement value given in L. Alton and Associates, *Czechoslovak National Income and Product (1947-1948 and 1955-1956)* (Columbia Univ. Press, 1962), p. 223.

Rumania: Growth rates of industrial output, capital stock, and labor were derived from Tables 36, 46, 47, and 49 in Directia Centrala de Statistica, *Anuarul Statistic al R.P.R., 1968*. Net capital stock and payments to labor used in estimating factor shares are from M. Montias, *Economic Development in Communist Rumania* (M.I.T. Press, 1967), pp. 57 and 58.

Yugoslavia: Growth rates of industrial output and labor were derived from Federal Institute for Statistics, *Statisticki Godisnjak, 1968*, pp. 164 and 170. Growth rate of net capital stock was derived from I. Vinski, "Fixed Assets, 1946-1966," *Yugoslav Survey*, Nov., 1968, p. 88. Net values of capital stock and payments to labor used in estimating factor shares are from I. Vinski, *ibid.*, p. 88, and Federal Institute for Statistics, *Inter-industry Relations for the Yugoslav Economy in 1962* (Belgrade, 1966), respectively.

Poland: Growth rate of industrial output derived from data provided by Thad Alton, Director, Research Project on National Income in East Central Europe. Growth rates of industrial capital and labor derived from Central Statistical Office, *Rosnek Statystyczny, 1968*, pp. 65, 100, and 180.

Hungary: Growth rates of output were derived from a series prepared by L. Czirjak of the Project on National Income in East Central Europe. Data on the capital stock, labor, and payments to labor were obtained from the *Hungarian Statistical Yearbook* and other publications of the Hungarian Statistical Office.

Greece: Growth rates of output and data on payments to labor and capital are from U.N., *Growth of World Industry*. Data on employment and the capital stock were made available by Professor Balopoulos, Director of the Center of Planning and Economic Research in Athens.

Ireland: Growth rates of output and data on payments to labor and capital are from U.N., *Growth of World Industry*, and manufacturing employment figures are from the International Labor Office, *Yearbook of*

Labor Statistics. Capital stock in manufacturing was estimated for 1853-1959 in Edward Nevin, *The Capital Stock of Irish Industry*, The Economic Research Institute, Paper No. 17. We made adjustments for the apparent overestimation of depreciation on investment prior to the second World War and extended the figures from 1959 on.

Spain: Growth rates of output and data on payments to labor and capital are from U.N., *Growth of World Industry*, and manufacturing employment figures from I.L.O., *Yearbook of Labor Statistics*. Estimates on the capital stock were made for 1965 in *La Riqueza Nacional de Espana*. These figures were adjusted on the basis of the gross investment figures provided in a private communication by Professor Lasuen of Madrid University and estimates of depreciation derived from national accounts statistics.

Norway: Growth rates of output and data on payments to labor and capital are from U.N., *Growth of World Industry*, and manufacturing employment figures from I.L.O., *Yearbook of Labor Statistics*. Capital stock data were made available by Odd Aukrust, Research Director of the Norwegian Central Bureau of Statistics.

REFERENCES

1. BELA BALASSA, "The Dynamic Efficiency of the Soviet Economy," *A.E.R.*, May, 1964, pp. 490-505.
2. ABRAM BERGSON, "National Income," in *Economic Trends in the Soviet Union*, A. Bergson and S. Kuznets, eds. (Harvard Univ. Press, 1963), pp. 1-37.
3. N. M. KAPLAN, "Retardation in Soviet Growth," *Rev. of Econ. and Statis.*, Aug., 1968, pp. 293-303.
4. R. M. SOLOW, "Investment and Technical Progress," *Mathematical Methods in the Social Sciences*, K. J. Arrow, S. Karlin, and P. Suppes, eds. (Stanford Univ. Press, 1960), pp. 89-104.
5. A. A. WALTERS, "A Note on Economies of Scale," *Rev. of Econ. and Statis.*, Nov., 1963, pp. 425-27.

DISCUSSION

JOHN M. MONTIAS: Anyone with the heart to read recent Czech or Slovak papers must agree with Professor Kým that the reforms, which still had some life left in them early this year, are moribund. The speeches of Prime Minister Černík and Minister of Planning Václav Hůla delivered before an assembly of Czechoslovakia's most important enterprise directors on October 17, 1969,¹ contained passages that could have been plucked from Antonín Novotný's final condemnation of Czechoslovakia's first set of reforms introduced in 1958-59, when he argued that central controls had been excessively loosened and that the only remedy was to strengthen the central direction of the economy. In contrast to 1963, however, the voices of dissent are now silent. Ota Šik is no longer here to contend, as he did on that earlier occasion, that the reforms failed "not because they went too far but because they did not go far enough."

In all fairness, though, it must be recognized that the strictly economic case for market reform in Czechoslovakia may not be so strong as it was a year or two ago. A foreign invasion, a balance-of-payments crisis, and domestic inflation hardly create favorable conditions for a smooth transition from a command to a market economy. An invasion followed by a political crisis cannot fail to shorten the planners' horizon and reduce the present value of the future gains in efficiency that are expected from the reform. Once foreign exchange and material reserves have been used up to meet immediate exigencies, planners averse to risk may see no alternative but to tighten the reins on domestic producers.

Let me turn now to a less melancholy subject. Portes' scholarly and dispassionate study of the Hungarian reforms comes to an optimistic conclusion which I am on the whole inclined to share, as long at least as the present politico-economic environment endures.

So far the most encouraging aspect of the Hungarian reforms has been the central authorities' ability to create relatively favorable macroeconomic conditions for decentralized decision making. They have avoided the trap into which the Novotný-led team of economic planners fell headlong in 1960-61 when they overloaded the investment program, stepped up output when bottleneck sectors were strained to capacity and precipitated a balance-of-payments crisis that helped wreck Czechoslovakia's first set of reforms. The seriousness with which the Hungarian authorities are taking the reforms is indicated by their willingness to cut back centralized investments to make room for an unanticipated rise in decentralized investments. They have used bank controls and fiscal policy (via the special measures for discouraging wage increases described by Portes) to ward off inflation, which, if not checked at the source, would invite the extension of price controls and other types of centripetal measures. (In Yugoslavia, for example, the efforts to combat inflation in the late 1950's led to temporary recentralization.) The subsidies the

Hungarians have continued to lavish on their export industries and the high tariff barriers they have imposed on imports, irrespective of their microeconomic justification, have helped to keep up foreign exchange holdings and to stave off a crisis which might have compelled the central authorities to intrude again into the detailed affairs of enterprises. A measure of macroeconomic stability, moreover, is a necessary condition for a further decontrol of producers' and consumers' prices, which in turn is essential for the reform to have the beneficial effects on efficiency that its designers are reckoning on.

Not that all the macroeconomic problems have been successfully resolved. A sellers' market still prevails in most sectors, but particularly in the capital goods industries. The fiscal instruments of the Hungarian reform described by Portes are still crude. A number of Hungarian economists, for instance, have pointed out the adverse effect on labor productivity of penalizing enterprises for increasing the average wage paid out to their workers. In Hungary's presently tight labor market, managers can hire workers at higher wages and still maintain the size of the enterprise's "sharing fund"—the principal source of their incentive payments—as long as the new workers are in relatively low skill categories and the weighted average of all wages is kept down. A rise in the wage bill relative to the value of the capital stock, moreover, may actually help to fatten the sharing fund because it allows the management to raise the sharing portion. These chinks in the reform are reported to have brought about the abnormal decline in the growth of labor productivity observed in the last eighteen months.²

Basing the development and sharing portions on the enterprise's capital-stock-to-wage-bill ratio has another potentially undesirable consequence: it discourages managers from taking any measures that might make their enterprise more capital intensive.³ To be attractive to managers seeking to maximize their sharing fund, an investment leading to an increase in the capital intensity of the enterprise must be much more profitable than in the absence of these profit-sharing rules. Rules restraining the managers' aggregate demand for investments may be a good thing, but there is no particular reason to expect them to induce an efficient discrimination between economically justified and unjustified projects.

This last consideration brings us to the microeconomics of the Hungarian reforms. Here much remains to be done before the system change begins to yield a clear improvement in resource allocation. Neither the price system nor the existing organization of enterprises, which effectively shields most of them from competitive pressures, can yet be counted on to support an efficient decentralization of decision making at the enterprise level. The main contribution of the reform so far has

¹ "Why Is Productivity Declining?" *Figyelő* (Budapest), No. 36, Sept. 3, 1969, p. 4, and "Income Regulation and Structure," *Figyelő*, No. 33, Aug. 13, 1969, p. 3.

² "Distribution Instead of Trade," *Figyelő*, No. 40, Oct. 1, 1969, p. 3.

¹ *Hospodářské noviny*, No. 43, Oct. 24, 1969.

been to free the enterprise of artificial constraints that formerly diminished its effectiveness. It is most probable, for instance, that useful technological and commercial information is now flowing much more freely throughout the economy, thanks to the possibilities now open to enterprises of engaging in spontaneous, mutually advantageous cooperation. Similar benefits may be expected from the closer contacts of domestic enterprises with their foreign suppliers or clients. This should eventually benefit Hungary's R and D efforts. To what extent inefficient prices will tend to distort the scarcity signals received by enterprises through these new channels and how these distortions will affect innovation and technical progress is not clear. Neither is the net effect of the change in system on risk-taking and innovative entrepreneurship, if one considers the entire range of decisions formerly arrogated by the center and now in part devolved upon enterprises.

We now come to what may be the fundamental problem inherent in the incremental transition from a centralized to a market system, where fiscal measures and monetary controls are meant to keep enterprises on the "right tracks" (at least until the market and competition have become sufficiently strong to exert the requisite pressures). If the center and its ministries are to guide the input and output decisions of enterprises through highly specific taxes and credit controls, they must collect and process as much information as if they were to issue the corresponding commands. (The detailed analysis of sales, costs, and profits is likely to rest ultimately on the same input coefficients as are used in material-balances planning.) If, on the other hand, these executive organs, in the spirit of reform, are bent on curtailing their operational responsibilities, then the coordination they ensure and the guidance they provide must be macro- rather than microeconomic in character. This leaves a greater scope for the enterprise to choose among its inputs and outputs and to mold its investment projects in accord with prevailing administered prices and other factors affecting profits, in the absence of any micro-coordination.

For these various reasons, I would anticipate a temporary drop in efficiency in the transition period of a major reform, at least if the former centralized system had been reasonably managed. Creeping recentralization to cope with manifest symptoms of inefficiency may easily occur in the transition, especially if the old ministerial apparatus has not been dismantled (as it was in Yugoslavia as early as 1951-52). If Portes is correct in his assessment, it augurs well for the Hungarian reforms that the ruling Party organs and the Council of Ministers are pushing ahead with the reforms and resisting the temptation of correcting its unsalutary side-effects by resorting to coercive measures.

Turning now to the economic model of workers' management that Jaroslav Vanek⁴ has just so eloquently defended, I am impelled to enter two serious reservations concerning the workers' choice between income

and security and the reaction of the system to exogenous changes in demand.

An important, though often neglected, condition for Pareto optimality is that every participant in a system should equalize the marginal rate of substitution between income and security, which may be defined, as a first approximation, as the inverse of the expected variance associated with his income. In a capitalist economy with a well-developed market for securities bearing various degrees of risk, every individual should be able to find the combination of employment and income-yielding assets most suitable to him, given the market value of his services and his initial wealth. In a collective economy of the Illyrian-type, however, an individual may not enjoy this opportunity. If he is extremely risk-averse, as most workers are likely to be since they have not enough wealth to take chances, the only option open to him is to seek admission to a collectively run enterprise with a record of steady profits where his income will be least subject to fluctuations.⁵ If he is located in a community where employment opportunities are limited or if his skills happen to be most useful in an industry, such as women's fashions, where sudden changes in demand can be expected to cause appreciable variation in profits and in workers' dividends, he will have no adequate way of protecting his income—unless he can somehow find some other individual(s) willing to reduce his risks in exchange for abandoning at least some portion of his claims to a fluctuating dividend. (Note, incidentally, that he will no more be able to buy insurance against the risk of low dividends, at least from profit-seeking insurers, than capitalist entrepreneurs wishing to insure themselves against low profits or nonrandom losses, as, in either system, the occurrence of losses in one firm is likely to be significantly correlated with the occurrence of losses in other firms in the same industry as well, possibly, as in firms belonging to other industries.) Only individuals more prone to take risks than he is can help the risk-averse members of the collective. But the exchange of security against income, whether or not it is carried out among men belonging to the same collective or to different firms, will subvert the system. A minority of the system's participants will end up taking most of the risks and engrossing most of the wealth. If individual adjustments of this sort were allowed to go on far and long enough, the Illyrian system would lose its distinguishing characteristics and become capitalistic in all but name. That such mutually advantageous exchanges are more than just a theoretical possibility is suggested by the experience of the Volkswagen Works with its stock-sharing plan: most of the workers to whom shares were distributed in 1960 are said to have sold them at the highest price they could get for them.

⁴ "Decentralization under Workers' Management: A Theoretical Appraisal," published in the Dec., 1969, *A.E.R.*, pp. 1006-14, but presented at this session.

⁵ If the collectively run enterprise, in order to cater to the risk aversion of most of its members, sets up a reserve fund to cushion the dividend from fluctuations in profits, then it freezes resources which could have been invested productively or consumed if decisions had been made by, or on behalf of, individuals willing and able to bear greater risks.

Few of the employees at the Renault plant, where a stock-sharing scheme has just been launched, are expected to hold on to them. At the European levels of real income, at any rate, workers still prefer cash to a share in the uncertain returns of the enterprise.

ment economy since then. The sequence of three different systems within a relatively short period of time offers a precious opportunity for an analysis of comparative efficiency of these systems. Some of the results are shown in the table below.

RATES OF GROWTH OF OUTPUT AND FACTOR INPUTS AND MEASURES OF FACTOR PRODUCTIVITY IN MANUFACTURING, MINING, CONSTRUCTION, AND CRAFTS IN YUGOSLAVIA, 1911-67

	Annual Rates of Growth			Total Factor Productivity
	Output	Capital	Labor	
Capitalism				
1911-32*	3.28	3.52	1.87	0.71
1932-40	4.67	2.59	0.72	3.16
1911-40	3.72	3.22	1.50	1.50
Central planning				
1940-54*	5.91	9.99	4.76	-1.04
Self-government				
1954-67	10.31	7.84	4.44	4.44

* War years 1914-18 and 1941-45 are left out.

SOURCE: B. Horvat, "Tehnički progres u Jugoslaviji," Institute of Economic Studies, Reprint No. 85, p. 48.

The adaptation of an industry made up of single-product firms to changes in demand, as Benjamin Ward has shown, would take place through the creation and the dissolution of firms rather than through the variation in the output level of existing firms. The adjustment process in Illyria would presumably be slower than in a capitalistic system, except in the limiting case where the firms facing an increase in demand in both systems were already operating at full capacity. If there were indivisibilities in the production process and/or delays in bringing new firms to the most advantageous scale of output, the industry's time lag in responding to market changes could easily lead to cobweb-type dynamic instability. Even if it were stable, the system's lags in response to demand changes would reduce the level of welfare that it could deliver, as compared to an otherwise identical system where the output of the individual enterprises was unconstrained.

BRANKO HORVAT: One may compare the growth performance of different countries in the same period of time. One may also compare growth and efficiency indicators of the same country at different points of time. The former approach was chosen by B. Balassa and T. Bertrand in the stimulating paper I have been asked to discuss. I propose to complement briefly their analysis by choosing the latter approach.

Yugoslavia is probably unique among all countries in having experienced all three basic contemporary socioeconomic systems within the time span of a generation. It passed through capitalist market economy before the war, it established centrally planned economy after the war, and it has been pioneering in self-govern-

ment economy since then. The sequence of three different systems within a relatively short period of time offers a precious opportunity for an analysis of comparative efficiency of these systems. Some of the results are shown in the table below.

$$\frac{\dot{Q}}{Q} = \gamma + \frac{\dot{L}}{L} + (1 - \alpha) \frac{\dot{K}}{K}$$

Labor coefficient $\alpha = 0.58$ is taken from the year 1967. Compared with capitalism, central planning increased the rates of capital and employment growth by three times. Since output growth lagged behind, the change in total factor productivity was negative. If Yugoslavia continued to practice central planning, it would, of course, become positive as in other centrally planned economy. The most likely interpretation of our data is as follows. The introduction of central planning speeds up growth, but increases inputs relatively more than output, and so lowers productivity as compared with other economic systems. Later on, total productivity of factors may grow at high rates but at any one time its absolute level is lower than in comparable economies in the other two systems.

The self-government period appears to be some sort of synthesis of the positive features of the preceding two periods. Employment growth continues to be almost as fast as under central planning. But output expands much faster, so that productivity growth achieves rates beyond anything known before.

	Rate of Growth of Output	Contributions to Output Growth in Percentage			
		Capital	Labor	Productivity Increase	Total
1911-40	3.72%	36.3	23.4	40.3	100.0
1954-67	10.31	31.9	25.0	43.1	100.0

The comparison of relative contributions to growth makes possible one additional conclusion. Since under widely differing conditions the relative contributions to growth appear to be quite stable in the Yugoslav economy, one may conclude that higher rates of output growth imply higher absolute increase in factor productivity. If inputs are interpreted as costs and productivity increase as "free gifts of nature," it follows that, unlike in classical economics, free gifts of nature may also be multiplied at will (up to a point, of course).

The last result corresponds to the hypothesis, tested by Balassa and Bertrand, that factor productivity depends also on the rate of growth of inputs. If in the formula

$$\gamma_2 = 1.60 + 0.45T_2$$

we insert my data for the capitalist and self-government period and if we compare the results computed with actual results, the outcome is as follows

$$1911-40: \hat{\gamma}_2 = 1.60 + 0.45 \times 2.10 = 2.55,$$

$$\text{actual } \gamma = 1.50$$

$$1954-67: \hat{\gamma}_2 = 1.60 + 0.45 \times 5.63 = 4.13,$$

$$\text{actual } \gamma = 4.44$$

The formula seems to be not too bad for high rates of growth, but fails to work for low rates. The relationship is obviously more complex and cannot be captured by a simple regression. But something else can be done.

It is possible to use data of Table 2 in the paper by Balassa and Bertrand to derive some interesting results shown in the next table.

By comparing capitalist, centrally planned and self-government economies we find the same regularities as when various stages of Yugoslav socioeconomic development were compared. The output rate of growth increases as we move from one group to the next. Employment expansion is particularly fast as we move from capitalism to central planning. (This, partly, explains the absence of overt unemployment in centrally planned economies.) The contribution of productivity to growth is least in centrally planned economies. Our initial hypothesis may be supplemented: not only the level of productivity but also its relative change is lower in centrally planned economies as compared with capitalist market economies.

I should end with a word of caution, however. The results cited are no more than empirical illustrations of some plausible hypotheses. Further research is required before more reliable conclusions can be established.

EVSEY D. DOMAR: My discussion of Professor Vanek's very stimulating paper runs into three difficulties: first, the subject is rather new and technical; second, his paper is essentially a summary of a large manuscript which I have seen but the audience has not; and, third, there is too little time to do justice to his paper, let alone his manuscript. I shall comment on a few points and let the others go.

The short-run market instability of an economy consisting of labor-managed enterprises, which I shall call here producer cooperatives, worries me less than it does Vanek. The alleged negative slope of commodity supply curves arises from the queer assumption that each co-op tries to maximize the dividend rate per unit of

RATES OF GROWTH OF OUTPUT AND FACTOR INPUTS AND MEASURES OF FACTOR PRODUCTIVITY IN MANUFACTURING INDUSTRIES IN COUNTRIES WITH DIFFERENT SOCIOECONOMIC SYSTEMS

	Annual Rates of Growth			Factor γ	Productivity for $\alpha = 65$ as % of the Output Growth Rate
	Output	Capital	Labor		
Capitalist economies*	7.1	6.3	2.5	3.3	46
Centrally planned economies†	8.7	8.1	4.1	3.0	35
Yugoslavia (self-government)	11.8	7.5	6.7	4.7	40

* Greece, Ireland, Norway, Spain.

† Bulgaria, Czechoslovakia, Hungary, Poland, Rumania.

NOTE: All quantities, except for Yugoslavia, are unweighted averages.

labor input. It is much more likely that the members and the management of a co-op will reach some understanding regarding the members' participation in work at each rate of dividend. Thus an increase in the price of output will call forth a larger labor effort rather than a smaller one, at least until the members work the desired hours of overtime. This assumes a positively sloping supply curve of labor—a safe assumption in an advanced country, though by no means in an underdeveloped one. The disagreement between Vanek and me on this point is much weaker in his paper than it was in his manuscript.

But where we disagree sharply is in our vision of the working of a co-op economy. Because of the inherent nature of labor management, Vanek expects each co-op to be small and reluctant to increase its membership. Thus an increase in the demand for some product will be met by an inelastic supply curve of existing co-ops (beyond the full-time employment of their members). The increased supply will have to come from new units. Their rapid emergence will also assure that no existing co-op can remain rich for long and pay higher-than-average dividends. Not only will this equality result in an equitable income distribution (equal pay for equal work), but also in the correct allocation of labor, since the dividend rate paid at or near the maximum point approximates the net value of the marginal product of labor. In the meantime, equal access to capital by the new co-ops will equalize its marginal rate of return as well.

Thus the principal equilibrating force in Vanek's system lies in the quick emergence of new co-ops. I believe that he thoroughly exaggerates the ease of organizing them. It is not only the question of constructing the capital, but even more so the difficulty of obtaining the skilled workers and the managers. These will have to be attracted (I assume the absence of compulsion) from successful co-ops which these persons will be reluctant to leave. In the meantime, the rich co-ops will remain rich, and may even become richer, since it will take a most unusual banker, even in a socialist state, not to favor loan applications from those who have proved their ability to succeed.

To some extent, dividend inequality among the co-ops can be reduced by taxation. But taxation must be used with care in order not to affect the members' incentive to perform.

I think it would be a great waste of proved ability not to encourage successful co-ops to expand by increasing their membership. And if the present members are reluctant to admit newcomers who would participate in the high dividends without having worked for them,

some seniority bonus could be arranged. (A nucleus of this idea was suggested to me by some visiting Czech economists before the Russian invasion; unfortunately, I do not remember their names.) This bonus might take the form of a few percentage points for each year of service applied to the number of man-days contributed by each member. It may be desirable to continue paying this bonus, or a good part of it, to a member after his departure from the co-op, and even after his retirement. This would increase labor mobility, particularly from the richer co-ops to the new ones, on whose emergence Vanek counts so much, and give each member a strong sense of security and of identification with his co-op.

While I find the basic idea of a democratic producer co-op very attractive, I cannot share Professor Vanek's enthusiastic, if not outright utopian, optimism. How does he know, for instance, that the competition caused by the co-ops' smaller size will at least compensate for the loss of economies of scale and of the reduction of risk in large units? How can he tell that the co-op members, in their egalitarian spirit, will not underpay talent and thus misuse it? The participation of all members may raise the wisdom of decisions, or it may reduce it to some average level. The members may welcome technological progress, but they may also find it too upsetting. They may work with greater devotion in a co-op, but they may also idle around in the absence of a boss. One sometimes has a feeling that Vanek refers not to a modern economy but to some medieval association of craftsmen or fishermen.

Obviously, our differences cannot be resolved by theoretical models. The actual outcome will depend most of all on the attitudes of the persons involved. An enthusiastic band of young volunteers, perhaps in Yugoslavia, may perform wonders in a co-op, such as a group of Russian peasants never will. It seems, then, that the freedom of participation is an important condition for the success of a co-op. But such freedom does exist in capitalist countries, and yet the scarcity of producer co-ops (except among some professional groups) is striking. Perhaps the spirit of capitalism is to be blamed for that.

But suppose all my doubts about the efficiency of co-ops come true, and a co-op economy, even if freely agreed upon, produces less or grows less rapidly than a capitalist one. Who can tell the participants that they are wrong if satisfaction from work and from participation in decisions is important to their welfare? Here I congratulate Professor Vanek for his emphasis on this point.

ECONOMICS OF EDUCATION

DEMAND AND SUPPLY IN U.S. HIGHER EDUCATION: A PROGRESS REPORT*

By R. RADNER, *University of California, Berkeley,*
and Churchill College, Cambridge
and

L. S. MILLER, *State University of New York, Stony Brook*

Introduction

Higher education in the United States may be thought of as a giant "industry," in which (1) the "inputs" are students of various qualifications, the services of teachers, and all the other usual kinds of material and service inputs, and (2) the "outputs" are graduates (and drop-outs) of various qualifications. Even the category of "goods in process" has its analogy in the case of students who are part way through a particular educational program.

If this picture is at all appropriate, then we shall be led naturally to study the demand for outputs, the supply of factors, the technological relationships within the industry, etc. Note, however, that the "supply of student inputs" from the point of view of the education industry corresponds to what is usually regarded as the "demand for places" by potential students. Colleges and universities do not sell in any direct way their output of graduates in the market for educated labor, nor are they usually thought of as selling their places to students. Furthermore, the hypothesis of profit maximization is no doubt even less appropriate to the education industry than to most conventional industries.

The present paper is a progress report on a project designed to estimate various supply, demand, cost, and technological relations in U.S. higher education. The project has essentially six component parts: (1a) estimation of student-teacher and other input-output relationships at the college and university level for a cross-section of institutions and (1b) estimation of student-teacher input-output relationships at the discipline or department level for a few selected public institutions in California; (2) a model relating dollar costs to various measures of activity; (3) estimation of the demand for places by potential

students as a function of cost of attendance, student family income, student ability, and school selectivity; (4) a study of the supply and pricing of places by private institutions; (5) estimation of the stocks of educated manpower, by age, sex, highest degree, and field of specialization, for a number of recent years; (6) a study of factors influencing the demand for educated manpower by the U.S. economy and thus for the output of the education industry. Of these, parts 1a, 3, and 5 are near completion, part 4 has only just begun, and the rest are at intermediate stages.

Our aim in this project is not only to add to the scientific description and understanding of the education industry but also to provide a set of related models that can contribute to the debate on policy issues. Effective policy analysis requires, of course, not just good projections based on the assumption of unchanging trends but also estimates of how policy instruments affect target variables. To illustrate the use of our models for policy purposes, we are applying them to several problems, including the estimations of (1) the resource requirements for universal two-year college attendance and (2) the effects of tuition increases and complementary financial aid programs in public institutions.

We should emphasize that there are a number of important topics with which this project does not deal, either because they are being intensively studied by others (e.g., the benefits of higher education, the supply of finance from sources other than tuition) or because we did not know how to tackle them with available data (e.g., the measurement of "quality" of inputs and outputs).¹ The main body of our paper is devoted to a description of some results from parts 1a and 3 of the project; i.e., on student-teacher input-output relations and on the demand for places.

* This project, which is being carried out at the Univ. of California, Berkeley, and S.U.N.Y., Stony Brook, is supported primarily by the Carnegie Commission on the Future of Higher Education, with supplemental support by the Office of Naval Research.

¹ For background papers on these and other topics, see *The Economics and Financing of Higher Education in the United States*, a compendium of papers submitted to the Joint Economic Committee of the Congress of the United States (U.S. Government Printing Office, Washington, D.C., 1969).

I. Teachers and Students

We first present some figures on the trends and dispersion of faculty-student ratios during the period 1952-64 for six different groups of institutions. We find, in particular, downward trends in this ratio in undergraduate colleges and in public universities, but upward trends in private universities (both nonsectarian and religious). The dispersion of the faculty-student ratios within groups declined in the undergraduate categories, but remained approximately stable in the university categories. Furthermore, there is a tendency for the ratio to decline most rapidly in schools with the highest ratio.

We then turn to a more detailed consideration of the relation between numbers of teachers and numbers of graduates and undergraduates, and various school characteristics such as faculty salaries, percent of faculty holding a Ph.D. degree, average SAT scores of entering freshmen, percent of students in teacher training programs, quality of graduate faculty, etc.

Faculty-Student Ratios in the "ACE Sample." Our first sample consists of 372 colleges and universities taken from a larger set of more than 900 institutions for which data were available² on numbers of faculty and students for the years 1952, 1956, 1960, and 1964. These 372 institutions included all those in the larger set that either were purely undergraduate institutions or had substantial graduate enrollment in each of the four years mentioned above but were neither purely graduate schools nor primarily religious or professional schools. Within the ACE sample these two groups will be called "undergraduate schools" and "universities," respectively; there are 259 undergraduate schools and 113 universities. For each of these institutions, and for almost every year, we have data on:

T = total faculty by highest degree attained,

S_u = number of undergraduate students enrolled,

S_g = number of graduate students enrolled.

After further subdividing the undergraduate schools and universities into the standard control categories of public, private-nonsectarian (hereafter called private), and private-sectarian (hereafter called sectarian), we calculated the average and the standard deviation of each of the resulting six groups for each of the four years in our observation period (1952-64). The results are presented in Table 1-A.

The mean faculty-student ratio clearly fell in each of the undergraduate groups, with the greatest decline (28 percent) in the public schools and the smallest decline (14 percent) in the private schools. The mean faculty-student ratio also fell slightly in the public universities but rose in the other universities. In both undergraduate schools and universities the private nonsectarian schools ended the period with the highest ratios and the public schools with the lowest; generally the private schools had the higher ratios throughout the period.

Of course, one suspects that the increases in the universities are due to the increased fraction of the total enrollment represented by graduate students. We shall have more to say on this later.

The variability of the faculty-student ratios, as well as their means, declined in the undergraduate school groups, but remained relatively constant in the university groups. We shall see below that, indeed, those undergraduate schools with the highest faculty-student ratios tended to suffer the most rapid decline. On the whole, there was considerable variation in the ratios, with the means roughly only two to four times the standard deviations. In 1964, the private universities had the lowest ratio of mean to standard deviation (1.9), whereas the sectarian undergraduate schools had the highest (4.5).

All in all, we have a picture of declining faculty-student ratios in undergraduate schools and in public universities and of increasing ratios in private sectarian and nonsectarian universities. The downward pressure on the faculty-student ratios seems most pronounced in the case of the public schools, both undergraduate and universities. Within each of the groups there is considerable variation in the faculty-student ratio. Our task will be to try to relate this to variation in institutional variables and, in the case of the universities, to changes in the undergraduate-graduate student mix.

The Relationship between Averages and Trends in the Faculty-Student Ratio. We have seen that there was a general decline in faculty-student ratios in the undergraduate schools between 1952 and 1964. To study this phenomenon in more detail, we measured, for each undergraduate school in our sample:

c_n , the average faculty-student ratio³ over the period 1952-64;

b_n , the average rate of change in the faculty-student ratio, per four-year period.

² American Council on Education, *American Universities and Colleges*, 6th through 9th editions, 1952, 1956, 1960, 1964.

³ For schools for which there were missing observations, the averages were computed for the available observations.

TABLE 1-A
AVERAGES AND STANDARD DEVIATIONS OF FACULTY-STUDENT RATIOS

	Undergraduate Schools			Universities		
	Public	Private Non-sectarian	Private Sectarian	Public	Private Non-sectarian	Private Sectarian
Mean of a						
1952	.0939	.1016	.1031	.0833	.1111	.0782
1956	.0859	.0963	.0973	.0785	.1244	.0868
1960	.0742	.0934	.0859	.0777	.1263	.0845
1964	.0674	.0875	.0809	.0766	.1266	.0956
Standard deviation of a						
1952	.036	.031	.033	.030	.062	.039
1956	.039	.028	.030	.029	.072	.037
1960	.028	.031	.023	.029	.074	.037
1964	.027	.022	.018	.030	.066	.041
Number of institutions in group						
1952	45	51	162	51	46	13
1956	45	51	162	54	46	14
1960	44	51	162	54	46	14
1964	45	51	162	55	45	14

$a = T/S$.

T = total faculty.

S = total students enrolled.

Table 1-B shows the mean and standard deviation of c_n and b_n in each of the three control categories. We see that the public schools had the lowest average ratio and the highest rate of decline, whereas the private schools had the highest average ratio and the lowest rate of decline.

However, an examination of the relationship between b_n and c_n on a school-by-school basis shows that the relationship by group is reversed. Table 1-C gives the regressions of b_n on c_n (rate on average) within each of the three groups and for the undergraduate schools as a whole. In each case the coefficient of c_n is negative (although statistically not significant for the private-nonsectarian schools). There is considerable variation around the regression lines, as the low values of R^2 indicate. Nevertheless, it is clear that there was a tendency for schools with higher average faculty-student ratios to decline more rapidly.⁴

Input Coefficients. It is generally believed that graduate students take up more faculty time, per

student enrolled, than do undergraduates. In the language of activity analysis, we might say that the training of undergraduate and graduate students are two different "activities," with different faculty input coefficients. This suggests the simple linear relationship:

$$(I.1) \quad T = a_u S_u + a_g S_g,$$

where, for a given school, at a given date, a_u and a_g are the faculty input coefficients for undergraduate and graduate teaching, respectively.

Direct estimation of equation (I.1), either from time series on individual institutions or from cross-sections of groups of institutions, has not produced satisfactory results. The data are simply not consistent with the hypothesis that, in general, for any one school the input coefficients are more stable than the ratio of undergraduates to graduate students. Nor have we yet found any convincing a priori classification of schools into groups with similar coefficient values.

The dispersion of faculty-student ratios among universities with the same graduate-undergraduate ratio is very large. For example, among private nonsectarian universities in which the

⁴ That this is not just another example of the "regression fallacy" is shown by Table 1-A which indicates that the variance of the faculty-student ratio is actually declining over the period.

TABLE 1-B
AVERAGES AND TRENDS OF FACULTY-STUDENT RATIOS FOR UNDERGRADUATE SCHOOLS

	Public	Private Nonsectarian	Private Sectarian	All
Mean c_n080	.095	.092	.090
Standard deviation c_n030	.025	.022	.025
Mean b_n	-.009	-.003	-.007	-.007
Standard deviation b_n011	.009	.010	.010
Number of schools.....	45	51	162	258

c_n = mean faculty-student ratio over the period 1952-64, for school n .

b_n = the average rate of change in the faculty-student ratio, per four years, for school n .

percentage of graduate students was roughly between 20 and 30 in the year 1966 (including, e.g., Adelphi University, Carnegie Institute of Technology, and the University of Rochester), the faculty-student ratio varied between .43 and .07 (i.e., the student-faculty ratio varied between 2.3:1 and 14:1).

We might suppose that the "crude" numerical input coefficients, a_u and a_v , depend upon the quality of the inputs and outputs, and possibly on other school characteristics as well.

Unfortunately, we have no accepted measures of the quality of inputs and outputs. However, it seems reasonable to suppose that schools with the same selectivity, tuition, faculty salaries, etc., will tend to have the same quality of inputs and outputs, or at least that the variation in quality among schools with similar characteristics is smaller than among schools with widely differing characteristics. This suggests that we try to estimate the relationship between the crude input-output coefficients and various school characteristics.

For each school, let W and Z be two vectors of measurements of various school characteristics (there may be some characteristics common to both vectors), and assume that the input coefficients depend upon these characteristics:

$$(I.2) \quad \begin{aligned} a_u &= h_0 + h \cdot W, \\ a_v &= k_0 + k \cdot Z, \end{aligned}$$

where h_0 and k_0 are parameters, h and k are vectors of parameters, and

$$h \cdot W = \sum h_j W_j, \quad k \cdot Z = \sum k_j Z_j.$$

Combining equations (I.1) and (I.2), and adding a constant term (c) yields, for each school, the equation:

$$(I.3) \quad T = c + (h_0 + h \cdot W)S_u + (k_0 + k \cdot Z)S_v.$$

The constant term, if different from zero, could reflect the presence of increasing or decreasing returns to scale. Equation (I.3) could also be applied to schools without graduate students, to examine how variation in the faculty-student

TABLE 1-C
REGRESSION OF b_n ON c_n : $b_n = \alpha + \beta c_n$

	Public	Private Nonsectarian	Private Sectarian	All
α , Constant term.....	.00185 (.42)	.00072 (.14)	.00138 (4.61)	.00677 (2.96)
β , Coefficient of c_n	-.131 (-2.54)	-.0437 (-.81)	-.226 (-7.16)	-.148 (-6.05)
R^213	.01	.24	.13
Number of observations..	45	51	162	258

NOTE: Numbers in parentheses are t statistics.

ratio is associated with corresponding variation in school characteristics.

We have estimated the parameters of equation (I.3) for different sets of data,⁶ different groupings of schools, and different sets of school characteristics. Space limitations do not permit a presentation here of the many regressions together with the many qualifications and reservations we would have to state concerning the specific formulations and numerical results. Nevertheless, the following qualitative relationships are suggested by the estimates.

There is a tendency for a higher percentage of faculty with the Ph.D. degree to be associated with lower undergraduate input coefficients, holding other variables constant. Since the other variables include average faculty salary for the school and since Ph.D.-holders typically command higher salaries than non-Ph.D.-holders, this would be a natural consequence of economizing with a given salary budget.

Our results thus far suggest a tendency in some cases for higher average faculty salary to be associated with higher input coefficients. If this is in fact the case, it indicates that "richer" schools use their additional funds both to pay higher salaries and to increase the faculty-student ratio.

There is also some tendency for universities with "higher quality" graduate schools to have higher graduate input coefficients. Here, a quality index was constructed by combining the school's ranking in the Cartter Report with a ranking by number of enrolled graduate NSF Fellowship holders. Finally, there is evidence for increasing returns to scale in the undergraduate institutions and for undergraduate teaching in the private universities.

These suggested relationships must be treated as tentative, pending further exploration of the data. Although certain institutional variables seem to have consistent effects on the input coefficients, we are not satisfied that we have quantitatively identified separate input coefficients for graduate and undergraduate teaching activities. One should be cautioned against extrapolating this effect beyond the range covered by our sample.

II. The Demand for Places

Our approach to the estimation of the demand for places in institutions of higher education has thus far focused on the decisions by individual

graduating high school seniors between going and not going on to college, and their choices among available institutions, or institution-types. We are dealing, therefore, with a demand more like that for houses or automobiles rather than for butter or beer, in the sense that the choice is among a small number of discrete alternatives rather than different quantities of a divisible good.

We imagine that each high school senior—hereafter called a "student"—faces a set of alternative choices. This set includes various types of institutions of higher education as well as the alternative of not going to any such institution. Our statistical model is designed to relate the relative frequencies of choices to the characteristics of the individual student and his alternatives. For actual estimation purposes we have available data for a sample of students included in the SCOPE study.⁷ The availability of data and the results of experiments with different formulations led us to concentrate on the following variables (whose precise definitions are given below):

- A_i an ability score for student i ;
- I_i a measure of income for student i ;
- S_j a measure of the "selectivity" or "quality" of alternative j ;
- C_{ij} the out-of-pocket dollar cost to i of going to j (set equal to zero for the alternative "no school").

We assume that the probability that student i chooses alternative j is a function of these variables, and the set of alternatives open to i , which we shall denote by J_i . We assume further that this functional relationship can be expressed in terms of two intermediary variables, to which (for the convenience of discussion) we have given the names "intellectual affinity" and "cost-to-income ratio" defined respectively by:

$$X_{ij} = \frac{A_i S_j}{1000}, \quad Y_{ij} = \frac{C_{ij}}{I_i}.$$

The particular functional relationship is a generalized form of logit analysis. For each i and j , define f_{ij} and P_{ij} by:

$$f_{ij} = aX_{ij} + bY_{ij}, \\ P_{ij} = e^{f_{ij}}$$

where a and b are parameters to be estimated. The conditional probability, P_{ik} , that student i chooses alternative k from the set J_i of alternatives open to him, given the values of the vari-

⁶ In addition to the ACE data already mentioned, we had access to 1966 HEGIS data made available through the Carnegie Commission on the Future of Higher Education.

⁷ SCOPE (School to College: Opportunities for Post-secondary Education), Center for Research and Development in Higher Education, Univ. of California, Berkeley.

ables X_{ij} and Y_{ij} , is assumed to be determined by the equation:

$$(II.1) \quad P_{ia} = \frac{F_{ia}}{\sum_{j \in J} F_{ij}}.$$

Note that this implies that the "odds" for any pair of alternatives, j and k , are equal to the ratio (F_{ij}/F_{ik}) , and the logarithm (to the base e) of these odds is equal to $(f_{ij} - f_{ik})$, or

$$a(X_{ij} - X_{ik}) + b(Y_{ij} - Y_{ik}).$$

The method of estimating the parameters a and b from data on a sample of students i is due to McFadden.⁷

We now turn to the definitions of the explanatory variables. "Ability" was measured by a test included in the SCOPE study, converted into "equivalent" SAT scores. "Selectivity" of an institution was measured by the average SAT score of entering freshmen, or an imputed average score for a category of "comparable" institutions (see remarks below on aggregation).

The cost C_{ij} was our estimate of the sum of tuition, living, and transportation costs, based on information that included knowledge of the locations of the institution and the student's home. In particular, we tried to take account of whether a given student could or could not be a commuter at a given institution. Unfortunately, we have thus far been unable to obtain data on the financial aid actually obtained by the students in our sample.

The measurement of income posed special problems. All of the students in the SCOPE study were asked to estimate their parents' annual income. For a subset of the students, there were income reports from the parents as well. The data showed very poor agreement between the figures reported by students and their parents. Therefore we experimented with estimating parent-reported income from student responses (available for all the students); these variables included parents' occupation, parents' education, parents' employment status, and the student's estimate of parental income. This resulting measure is called here "predicted income," whereas the figure given by the parents (when available) is called "reported income." Within the sample of "reporters," the income prediction equation explained

45 percent of the variance in parent-reported income.⁸

We also experimented with a measure of income that might be more related to "ability to pay." For this purpose we chose the concept of "discretionary income" used by the College Scholarship Service in evaluating need for financial aid.⁹ This measure is a function of a family's net income before taxes and the number of dependent children.

Combining the two dichotomies, reported versus predicted and total versus discretionary, we had four different measures of income with which to experiment.

For the results given in the present paper, our sample included approximately 190 students in each of two states, California and Illinois. The sample from each state was itself made up of two subsamples, corresponding to whether the parents did or did not report income. The parameters a and b of the conditional probability function were estimated separately for each subsample and for each measure of income available for that subsample (yielding six sets of estimates for each state).

In order to make the data collection and the estimation feasible with our limited resources, we aggregated institutions into "school types"; Table 2-C shows the types that we used and their typical characteristics. Finally, for each student in the sample, our estimate of the set J_i of school types for which he was eligible was based upon his ability score, the selectivity of the school type, and, in the case of public institutions, a direct verification of the individual's eligibility.

Tables 2-A and 2-B give the estimates for California and Illinois, respectively. We note first that in all cases the estimate of b , the coefficient of the "cost-to-income ratio" is negative (as we would expect) and significantly different from zero (statistically). Furthermore, between states there is reasonably good agreement between corresponding estimates for the two subsamples ("parent-reporters" and "parent nonreporters"), using of course only the predicted measures of income. In each of these four cases the b coefficient

⁸ For a more detailed discussion of our procedure and problems see L. S. Miller, "Predicting Family Income in the SCOPE Sample," Carnegie Commission Project Working Paper, Dept. of Economics, Univ. of California, Berkeley.

⁹ See *Manual for Financial Aid Officers*, College Scholarship Service, 1967, Chap. 5 and Appendix B. Our conversion of total income to discretionary income is based on tables in J. E. Nelson, *Student Financial Aid Administration, Requirements, and Resources at the University of California*, Part II, pp. 7-9, and Appendix F, pp. 4-5.

⁷ D. McFadden, "The Revealed Preferences of a Government Bureaucracy," Technical Report No. 17, Project for the Evaluation and Optimization of Economic Growth, Dept. of Economics, Univ. of California, Berkeley, Nov. 1968 (mimeographed).

TABLE 2-A
DEMAND FOR FRESHMAN PLACES IN HIGHER EDUCATION INSTITUTIONS:
1966 CALIFORNIA HIGH SCHOOL SENIORS

Sample	Problem	(Ability \times Selectivity)/1000	Cost/Predicted Income	Cost/Discretionary Predicted Income	Cost/Reported Income	Cost/Discretionary Reported Income
I	1	$-.321 \times 10^{-3}$ (.586 $\times 10^{-3}$)	-11.39 (2.45)			
	2	-0.848×10^{-3} (.589 $\times 10^{-3}$)		-.848 (.218)		
II	1	$.129 \times 10^{-3}$ (.532 $\times 10^{-3}$)	-9.77 (1.73)			
	2	$.738 \times 10^{-3}$ (.523 $\times 10^{-3}$)		-.832 (.176)		
	3	$.113 \times 10^{-3}$ (.507 $\times 10^{-3}$)			-3.74 (1.10)	
	4	$.565 \times 10^{-3}$ (.478 $\times 10^{-3}$)				-.689 (.191)

SAMPLE I: Sample of students whose parents did not respond to SCOPE parent questionnaire (96 observations).

SAMPLE II: Sample of students whose parents did respond to SCOPE parent questionnaire (96 observations).

Figures in parentheses are standard deviations of the corresponding estimates.

for the parent nonreporters is somewhat larger in magnitude than the corresponding coefficient for parent reporters, indicating that the choices of the nonreporter students were more sensitive to cost differences (relative to their family incomes) than were the choices of reporter students.

Within the sample of reporters in each state, if we use discretionary income, then the estimates of b are about the same whether we use predicted or parent-reported income. However, the use of total income leads to estimates that are about twice as large for the case of predicted income as for the case of parent-reported income.

The "intellectual affinity" variable does not turn out to be as significant as we had expected. In the three cases in which a was significant, it was estimated to have similar positive values. It

is interesting to note that these a 's were always found in conjunction with the predicted measure of income. Indeed, only one of our six specifications had both variables significant in both states: Sample II, Problem 1.

Thus, our most satisfactory estimates in both states were for the reporter samples, where predicted income rather than parent-reported income was used to compute the variable "cost-to-income." If we look at the variables that enter the income-prediction equation (see above), we see that predicted income might well be a good index of social and educational status, even though it was originally designed purely for income prediction purposes.

It should be remembered that our cost variable does not reflect the actual financial aid received

TABLE 2-B
DEMAND FOR FRESHMAN PLACES IN HIGHER EDUCATION INSTITUTIONS:
1966 ILLINOIS HIGH SCHOOL SENIORS

Sample	Problem	(Ability \times Selectivity)/1000	Cost/Predicted Income	Cost/Discretionary Predicted Income	Cost/Reported Income	Cost/Discretionary Reported Income
I	1	$.165 \times 10^{-1}$ ($.582 \times 10^{-2}$)	-13.32 (2.04)			
	2	$.107 \times 10^{-1}$ ($.572 \times 10^{-2}$)		-1.051 (.193)		
II	1	$.103 \times 10^{-1}$ ($.545 \times 10^{-2}$)	-9.33 (1.61)			
	2	$.337 \times 10^{-2}$ ($.538 \times 10^{-2}$)		-.560 (.128)		
	3	$.448 \times 10^{-2}$ ($.526 \times 10^{-2}$)			-4.58 (1.19)	
	4	$-.650 \times 10^{-2}$ ($.499 \times 10^{-2}$)				-.536 (.137)

SAMPLE I: Sample of students whose parents did not respond to SCOPE parent questionnaire (91 observations).

SAMPLE II: Sample of students whose parents did respond to SCOPE parent questionnaire (90 observations).

Figures in parentheses are standard deviations of the corresponding estimates.

by the student; therefore, estimates of b are probably biased downward in absolute value and our estimates of a may also be biased to the extent that financial aid is correlated with ability.

It may help the reader to interpret these results if we give a few illustrations of predicted choice probabilities, based upon California Sample II, Problem 1. Here the estimates of a and b are 0.129×10^{-1} and -9.78 , respectively, and we use predicted income. Consider a high school senior from a family with a \$6,000 annual income, whose ability corresponds to an average SAT score of 550, and who lives in an area where he could attend a public junior college, state college, or university and still live at home. He has open to him a range of 11 choices, including "no school." These choices, together with the corre-

sponding predicted probabilities, are shown in Table 2-C. For example, the probability of "no school" is predicted to be .35, and the probability of choosing the local campus of the public university is .15. Table 2-C also shows choice probabilities for students from families with \$12,000 income and ability score 450. Note that the ranges of choice are not the same in all cases.

Equation (II.1) can also be used to calculate the changes in the choice probabilities that would be associated with a change in one or more of the explanatory variables. For example, for a student with income \$6,000 and ability score 650, an increase in the cost of going to the local public university of \$100 would decrease the predicted probability of that choice by approximately .037, which is more than a 10 percent decrease in the

TABLE 2-C

PREDICTED PROBABILITIES OF CHOICES FOR SELECTED CALIFORNIA STUDENT TYPES

Income of i Ability of i School type k			\$6,000 550	\$12,000 550	\$6,000 650	\$12,000 450
	Cost C_k	Selectivity S_k				
No school.....	0	374	.35	.22	.35	.30
Public junior college, local.....	290	430	.21	.18	.22	.24
Private medium cost junior college.....	2,200	430	.01	.04	.01	.05
Private high-cost junior college.....	3,200	430	.00	.02	.00	.02
Public state college, local.....	400	519	.18	.17	.18	.22
Public state college, not local.....	1,300	519	.04	.08	.04	.10
Private medium cost college.....	2,200	519	.01	.04	.01	.05
Private high-cost college.....	3,200	519	.00	.02	.00	.02
Public university, local.....	540	564	.15	.15	.15	
Public university, not local.....	1,440	564	.03	.07	.03	
Private university,						
high cost.....	3,200	564	.00	.02	.00	
"Superior" private uni- versity, high cost.....	3,200	625			.00	

probability (this is based on the above parameter estimates).

In our further work, we shall look at student

samples from two other states, Massachusetts and North Carolina, and experiment with refinements of our various measurements for all four states.

INCOME DISTRIBUTION EFFECTS OF HIGHER EDUCATION*

By W. LEE HANSEN
University of Wisconsin

Whether—and if so, how—to increase the resources available to finance higher education is a question raised ever more frequently by educators, legislators, and the general public. The vast enrollment growth of the early and middle 1960's, combined with sharp price increases since then, has produced a large and continuing demand for additional resources. At the same time the revenue-generating abilities of institutions of higher education appear not to have kept pace, largely because of other rapidly growing demands on public funds. The result has been a gradually tightening financial squeeze, accentuated by recent federal cutbacks in funds and by reduced generosity on the part of state legislators responding to campus turmoil.

Partly in anticipation of this squeeze and partly in response to it, various proposals have been advanced that would increase the resources available for higher education. Among these proposals are an expanded program of state and/or federal institutional grants, income tax credits, contingent repayment student loan programs, larger amounts of conventional student financial aid, and the like [1] [6]. While all of these proposals would directly or indirectly increase the resources devoted to higher education, they would provide these resources in different ways, to different people, and with varying effects.

To evaluate the desirability of these proposals requires that we know how the present system of financing higher education operates to promote the objectives of economic efficiency and equity. This paper explores one part of this much larger topic, by providing new information on the equity or income distribution effects of higher education. The emphasis will be on the tax-supported systems which provide substantial subsidies to young people and the parents of young people enrolled in public colleges and universities. The focus will necessarily be on the state level, since

the amounts of these subsidies, who receives them, and who pays for them are largely the result of state rather than federal policy. Moreover, because the income distribution effects are likely to differ from state to state, we examine the results for two different states, California and Wisconsin. This requires estimating the nature and magnitude of these distributive effects for Wisconsin, and then comparing them with a similar study for California. To keep the discussion within manageable bounds, attention is confined to undergraduate education.

I

An assessment of the distributional impact of higher education is important for several reasons. First, it fills an important gap in our knowledge about the distributional effects of a large and important public program operating at the state level.¹ Second, in view of the oft-stated goal of greater equality of opportunity, such an assessment makes it possible to examine the extent to which this goal is being realized. Finally, firmer knowledge about the magnitude of income distribution effects will be helpful in evaluating both the efficiency and equity effects of alternative proposals for the financing of higher education.

We already possess some knowledge of the distributional effects of public higher education from a recently completed study for California [3]. First, taxpayers in general subsidize the families of young people enrolled in public institutions of higher education. Second, larger subsidies go to those families whose children are eligible for and enroll in higher-cost, higher-prestige institutions; and it is these families who on average have higher incomes and are most able to pay. Third, these higher-income families do not pay commensurately higher state and local taxes in California, and so on average there is for them a higher ratio of education subsidies to state and local taxes paid.

These results hinge upon several key parameters: the structure of the tax system which provides funds for the support of higher education, the family income distribution of students

*I wish to acknowledge the useful comments of Burton A. Weisbrod, the helpfulness of Billy Dee Cook, S. E. Horkan, and David R. Witmer in supplying unpublished data, and the research assistance of Ann S. Jondrow. The research reported here was supported in part by funds granted to the Institute for Research on Poverty at the University of Wisconsin by the Office of Economic Opportunity pursuant to the provisions of the Economic Opportunity Act of 1964. The conclusions are the sole responsibility of the author.

¹ The need to examine the redistributive effects of government programs, in addition to the efficiency effects, has been emphasized by both Burton A. Weisbrod and James T. Bonnen [4].

enrolled in different schools, tuition charges, and the level of full educational costs per student at different schools. Because these parameters are likely to vary from state to state, it is important to replicate the California study so that the broader pattern of income distribution effects will emerge more clearly. For example, we know that in California admission to different types of colleges depends upon high school performance, with the standards being highest at the University of California (UC), lower at the state colleges (SC), and lowest at the junior colleges (JC). We also know that the subsidy received by a student is greatest at the UC, somewhat smaller at the SC, and smallest at the JC. Third, we know that admission standards give rise to different types of student clientele at each of the three systems, with on average the UC having the highest-income students and the JC having the lowest-income students. Finally, we know that because of its state income tax, California's overall tax structure is less regressive than that of most other states.

Many states would serve as good candidates for comparison with California, but lack of data, particularly on the incidence of taxes by family income level, greatly reduces the number of states that can be considered. Indeed, the only state for which reasonably good data are now available is Wisconsin. Since Wisconsin, like California, relies partly on an income tax, it would be most interesting to compare still another state which relies heavily upon sales taxes as a source of revenue.²

We turn now to estimate the income distribution effects of higher education in Wisconsin. The value of the subsidy available to a Wisconsin resident who is a student in a public institution in Wisconsin is the difference between tuition and the full costs of college education.³ The costs of college are taken to include not only instructional costs but also operational and capital costs.⁴

² A recent study for Florida indicates that public higher education redistributes income from lower- to higher-income groups. Since the data and methodology of the study differ, the Florida results cannot be compared with those here [7]. Another study indicates that the income distribution effects of higher education in Canada are negligible [5]. For critical comments on the Hansen-Weisbrod study [3] and a reply, see recent and forthcoming issues of the *J. of Human Resources*.

³ Throughout the paper, we are concerned with "in-state" students—Wisconsin residents who were paying about 20 percent of instructional costs. "Out-of-state" students were paying approximately 100 percent of instructional costs.

⁴ For instructional and operating costs we rely upon the traditional college accounting data. The cost of the services provided by capital (buildings and equipment) are far more difficult to estimate with precision; rough estimates have been developed for use here.

Despite differences in the apparent "quality" of the University of Wisconsin (UW) system and the Wisconsin state university (WSU) system, the full institutional costs in 1964-65 were approximately equal, at \$1,200 per academic year. Since tuition amounted to \$300 at the UW and to \$190 at the WSU,⁵ net per-student institutional costs—or the subsidies per student—amounted to \$900 and \$1,010, respectively. The effect of this tuition differential is to provide larger net subsidies to WSU students and their families than to UW students and their families.

The ability to pay of families and students differs in the two systems, as indicated by the median family income levels. Families of UW students report median incomes of \$9,700 per year in 1964-65, considerably higher than the \$6,500 estimated family income of WSU students. Because Wisconsin's state tax system is progressive over this income range, the tax contribution differs considerably for the median families with students in these two systems.⁶

Information on the single-year subsidies, family income levels, and state taxes can now be put together to show the distribution effects of Wisconsin's publicly supported system of higher education. Lines 1 to 5 in Table 1 are self-explanatory. Line 6, labeled "Net transfers", refers to the difference between the subsidy received and total state taxes paid in that year. It is important to remember that Wisconsin taxes go to defray the costs of a wide array of state-provided services, including some revenue-sharing with local communities. Hence, the state taxes included here do not reflect the taxpayer contribution to higher education alone; unfortunately, there is no easy or correct way to determine what portion of a family's taxes is used to provide any particular service, such as higher education. At the same time it is clear that taxes are generally paid over many years while college subsidies are received only while the student remains in college.

Several interesting results emerge from Table 1.

⁵ Actually, the "tuition" figures include both "tuition and fees." Since tuition (payment for instructional costs) is less than total tuition and fees, the subsidies are slightly understated. The impact of state scholarships, based largely on financial need, cannot be estimated because of the lack of adequate data.

⁶ The state tax structure in Wisconsin is more progressive than in most states. The marginal tax rate rises from 4 percent at the \$3,000-4,000 level of net taxable income to 8 percent at the \$14,000 level of net taxable income, and to a 10 percent maximum. However, the marginal rates based on gross income, the nearest equivalent to family income, are considerably lower and less progressive, rising from about 4 percent at the \$3,000 level to about 6 percent at the \$14,000 level.

TABLE 1

AVERAGE FAMILY INCOMES, AVERAGE HIGHER EDUCATION SUBSIDIES RECEIVED,
AND AVERAGE STATE TAXES PAID BY FAMILIES, BY TYPE OF INSTITUTION
CHILDREN ATTEND IN WISCONSIN, 1964-65

	All Families	Families without Children in Wisconsin Public Higher Education	Families with Children Enrolled in Wisconsin Public Higher Education	
			University of Wisconsin	Wisconsin State Universities
	(1)	(2)	(3)	(4)
1. Average family income.....	6,800	6,500	9,700	6,500
2. Average higher education subsidy per year.....	—	0	900	1,010
3. Average subsidy as a percent of family income.....	—	0	9.3	15.8
4. Average state taxes paid.....	240	240	430	270
5. Average state taxes paid as a percent of family income.....	3.6	3.6	4.3	4.1
6. Net transfers (line 2—line 4) ..	—	—240	+470	+740

SOURCE: Based on unpublished information from University of Wisconsin, Wisconsin state universities, Wisconsin Department of Administration, and Wisconsin Department of Revenue, and on estimates by the author.

First, families with children in college are subsidized—at least temporarily—by families who do not have children in college; this includes some families with children of college age not in college. Second, the largest annual subsidies go to WSU students, as already noted; in addition, as a percentage of family income these subsidies are even more favorable to WSU students compared to UW students. Third, the absolute amounts of the net transfers are higher for WSU students than for the generally more affluent UW students.⁷

How do these results compare with those for California? To answer this question, several adjustments must be made in the California data [3, Chap. 4]. One requires that the junior college system be excluded because, as of 1964-65, Wisconsin had no comparable two-year college system.⁸ This adjustment necessitates another one: the exclusion of local taxes which in California

provide substantial support for the JC system.⁹ Thus, our comparison is limited to the two major systems in each state and to those fully dependent upon state financing.

The contrasts between the two states are sharp, as revealed by a comparison of Tables 1 and 2. Subsidies per student are substantially higher in California, in part because of lower tuition (essentially zero at that time) and in part because of larger expenditures per student. More interesting is the finding that subsidies tend to be proportional to family income of students in the two California systems (Table 2, row 3) but are inversely related to family income for students in the two Wisconsin systems. Net transfers—which reflect the structure of state taxes as well as subsidies—are also proportional to family income in California. In Wisconsin, however, the higher level of taxes paid and the wider differences in family incomes of students combine to make net transfers relatively more redistributive be-

⁷ Were the state tax structure more regressive, through greater reliance on sales taxes, then net transfers would be less redistributive for WSU students and their families vis-à-vis UW students. However, the redistribution from the noncollege population to the college-going population would be even greater than it is now.

⁸ With the full emergence of the vocational-technical college system and with the availability of new data, it should soon be possible to update the Wisconsin results, and to broaden them to include all postsecondary undergraduate education.

⁹ Despite this adjustment, the comparisons still leave something to be desired on the tax side. For example, Wisconsin uses its taxing power to share revenue with local units; while this reduces the local property tax rate, it increases the state rate. In addition, both states provide some support for their third-level systems—JC's in California and vocational-technical schools in Wisconsin.

TABLE 2

AVERAGE FAMILY INCOMES, AVERAGE HIGHER EDUCATION SUBSIDIES RECEIVED,
AND AVERAGE STATE TAXES PAID BY FAMILIES, BY TYPE OF INSTITUTION
CHILDREN ATTEND IN CALIFORNIA, 1964-65

	All Families	Families without Children in California Public Higher Education	Families with Children Enrolled in California Public Higher Education	
			University of California	California State College
1. Average family income.	(1) 8,000	(2) 7,900	(3) 12,000	(4) 10,000
2. Average higher education subsidy per year.	—	0	1,700	1,400
3. Average subsidy as a percent of family income.	—	0	14.2	14.0
4. Average state taxes paid.	192	182	350	260
5. Average state taxes paid as a percent of family income.	2.4	2.4	2.9	2.6
6. Net transfers (line 2—line 4). .	—	—180	+1,350	+1,140

SOURCE: Adapted from Table IV-12, p. 76, in W. Lee Hansen and Burton A. Weisbrod,
Benefits, Costs, and Finance of Public Higher Education.

tween the two systems than are net subsidies alone. Hence, the Wisconsin system appears to be termed more egalitarian than the California system for those people attending college.

II

If we now broaden our view to consider all public postsecondary schooling, we would find that both states provide substantial subsidies to well over half of their recent high school graduates—those going to public colleges. Many of these students, moreover, avail themselves of these subsidies for more than a single year of college; and in general the higher the family income the more likely the student is to complete college. But, no subsidy whatsoever goes to the rest of the high school graduates who do not attend public colleges. Many of these young people come from families with modest incomes. Whether they fail to attend college because they are not accepted for admission, because they are not encouraged to enroll in college, or because they simply prefer not to attend, the fact remains that they receive no subsidy at all. Neither, of course, do the more affluent who prefer enrolling in private schools.

The existence of this pattern of income redistribution through the public expenditure system, as well as through the tax system, has long been suspected even if not fully documented. Paradoxically, the most common justification for the re-

distribution in the case of higher education is that it helps to achieve greater equality of opportunity. The lower the tuition, the easier it is for students to attend. But low or even zero tuition has still not been sufficient to permit sizable numbers of young people, particularly from lower-income families, to attend public institutions. What has happened is that low tuition provides a large subsidy that is given out indiscriminately to every enrolled student, on the grounds that anyone enrolling is deserving of a subsidy. But when public funds for subsidies are limited, as they inevitably are, the proper question is: Who needs them the most? By and large, the need is greater for qualified students from lower-income families, many of whom now either do not go to college at all or if they do go, are more likely to drop out early and/or to incur substantial debt while in school.¹⁰

Another part of the justification for low tuition and the income redistribution which it promotes hinges on the external benefits that are presumed to result from higher education. The reasoning is that because these benefits accrue to society as a whole, society should in effect compensate through general tax support those going to college. But this appealing and frequently-heard argument stems from a faith in rather than firm knowledge of the existence of and possible magnitude of these

¹⁰ See [2] for a fuller discussion of the low-tuition argument.

external benefits. While the research activities of colleges and universities, for example, may produce sizable external benefits, it is much more difficult to make a similar case for undergraduate education. I myself am skeptical about the external benefits justification. Until better evidence on these benefits is forthcoming, this justification for general tax support continues to be a weak one.

From another point of view, it may be argued that, if a longer time period is considered, little or no actual subsidization of college students occurs. Because higher education leads to higher incomes, it is argued, students will in later life pay substantially more tax revenue to the state and in its taxpayers, in effect, repaying the value of the subsidies received during their college years. But whether the amount expected to be repaid is sufficient to offset the subsidies received is an empirical question.¹¹ The present value of additional state and local taxes expected to be paid in California falls considerably short of the present value of the subsidy received by either male college graduates or those males completing only two years of college [3, Chaps. 2 and 4]. The gap between future tax payments and subsidies would be even greater if the calculations were repeated taking account of state taxes only. For Wisconsin it also appears that the subsidies would be only partially offset by the additional future tax payments to the state.¹² What accounts for the gap is that the combined effect of the additional taxable income associated with (or resulting from) college going, the level of state tax rates, and the progressivity of the tax structure is not great enough to produce a sufficient lifetime increment to tax revenue. In any case, the gap would be expected to be smaller in Wisconsin than in California, because the former's subsidies are smaller and the effective state tax rates are higher. The difficulty in recouping past public subsidies is compounded, however, because considerable numbers of young people who benefit from higher education subsidies migrate from the state and in this way escape all or at least a part of the repayment via taxes.

In our focus on the public sector, there has been no discussion about the equity of a system whereby parents of students attending private colleges and universities contribute tax revenues which help to support public higher education even though these families receive no direct sub-

sidies themselves. Concern about this group has been somewhat less widespread, largely, it would seem, because smaller numbers of people are involved—only about 10 percent of the actual college-going population. Yet application of the low tuition and tax repayment arguments would, in the interests of symmetry, also call for the subsidization of students in private higher education. In the absence of such subsidization, a redistributive effect occurs between families with children in public and those with children in private colleges. But the extent of concern about this redistribution is lessened by two factors: private school students tend to be from wealthier families who can more easily afford to pay the costs of higher education; in addition, a conscious decision has been made not to attend a public institution offering a subsidy.¹³

The equity of a system of restricted subsidies to college-going young people has received little attention. Such a system seems to assume implicitly that college going is the primary, if not the sole, means of enhancing potential earning power and/or the prospects for a satisfying, enriched life. But at least roughly similar beneficial effects seem likely to result from other types of education and training programs, among them technical training courses, conservatory programs, apprenticeships, on-the-job training, and the like.

What the existing subsidy system does is to encourage individuals to invest in higher education by making higher education relatively inexpensive. Meanwhile, young people who may recognize the inappropriateness of college to their own vocational aspirations are discouraged from pursuing alternative programs because they must pay the full (unsubsidized) costs of these programs. Yet these young people and their parents, who on average are less able to pay, continue to be taxed to support the college training of others.

In the interests of promoting greater equality of opportunity as well as widening the options open to young people, eligibility for public subsidies might be broadened to include other types of education and training in addition to college. Ample precedent exists in the GI Bill and manpower training programs for enlarging the range of programs in which students can be subsidized.

¹¹ Also at issue is the extent to which additional schooling can be associated with additional earnings, although the evidence seems to be reasonably favorable on this point.

¹² Based on calculations made by the author.

¹³ In 1965-66 Wisconsin instituted a subsidy program to resident students enrolled in private colleges in the state. A maximum grant of \$500 per year was made available, to be based on family ability to pay. Families with "effective incomes" of less than \$2,000 qualified for the full \$500; the subsidy declined at a steady rate such that at an effective income of \$10,000 no subsidy was received. The program has been raised somewhat since then.

If a broadened subsidy program is to be considered, however, we must again confront the question of financing—whether through additional public funds or through a reduction of the subsidies now going to the college-going population.

III

This paper attempts to indicate the nature of the income redistribution effects of the public financing of higher education. The purpose has been to develop a base against which to compare the effects of alternative proposals for the financing of higher education. How the different proposals stack up is beyond the scope of this paper. What seems clear, however, is that there are important and not insubstantial distribution effects.

Whether society wants to continue to produce these redistribution effects through its current methods of financing higher education remains to be answered.

REFERENCES

1. CARNEGIE COMMISSION ON HIGHER EDUCATION, *Quality and Equality: New Levels of Federal Responsibility for Higher Education* (McGraw-Hill, Dec., 1968).
2. W. LEE HANSEN AND BURTON A. WEISBROD, "A New Approach to Higher Education Finance," Discussion Paper 64-70, Inst. for Res. on Poverty, Univ. of Wisconsin, Feb., 1970.
3. ———, *Benefits, Costs, and Finance of Public Higher Education* (Markham, 1969).
4. JOINT ECONOMIC COMMITTEE, *An Analysis and Evaluation of Public Expenditures: The PPB System*, Vol. I (U.S. Government Printing Office, 1969).
5. RICHARD W. JUDY, "On the Income Redistributive Effects of Public Aid to Higher Education in Canada," Institute for Quantitative Analysis of Social and Economic Policy, Univ. of Toronto, Sept., 1969.
6. U. S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE, *Toward a Long-Range Plan for Federal Financial Support for Higher Education: A Report to the President* (U.S. Government Printing Office, Jan., 1969).
7. DOUGLAS M. WINDHAM, "State Financed Higher Education and the Distribution of Income in Florida" (unpublished Ph.D. dissertation, Florida State Univ., June, 1969).

RESOURCE ALLOCATION IN HIGHER EDUCATION

By JOSEPH A. KERSHAW, *Ford Foundation*

and

ALEX M. MOOD, *University of California, Irvine*

When we were asked last summer to put this paper together for the Christmas meetings, it was our hope that we would have some impressive new developments to report on the way higher education decides how to allocate its resources. As it turns out, all we can do is indicate some of the kinds of research and experimentation that are now in progress and some of the problems being faced. The Ford Foundation has made a number of substantial grants to colleges and universities that are experimenting with new ways of budgeting, of analyzing the decision-making process, and of constructing models to help the decision-maker understand the consequences of his decisions. The Public Policy Research Organization at Irvine is involved in the taking of a critical inventory of what we know and do not know in this elusive area and of what progress we are making. The final report of the Irvine project is one year away, though, and has little to contribute here. That study, incidentally, is being funded jointly by the Ford Foundation and the Carnegie Commission for Higher Education, the latter chaired by our chairman of this session.

Economists need not be told of the importance of intelligent decision making in higher education. Society's annual expenditure of resources on higher education is rapidly approaching the \$20 billion mark and the total cost to students (including opportunity costs) is of the same order of magnitude; the rate of growth of these sums shows as yet no sign of slowing down, though the rate will slow of course as the proportion of those of college age in college approaches its maximum. In any case these costs are large and will continue to be large. If we can improve utilization of those resources by 10 or 20 percent, the benefit to the nation will be very significant. Perhaps as important, the suppliers of those resources—taxpayers, alumni, parents, foundations—who are called on each year to increase their generosity have a right to know that the resources they supply are being used effectively. It is incumbent upon higher education, which above all institutions should be characterized by an insistence on a rational process, to provide assurance that it is consuming resources effectively.

What is new in the way people have been looking at the higher educational problem in

recent years is the attempt to treat it as a system, whether we are talking about individual institutions or all institutions. The parentage of this new way of looking at higher education is clearly the program budgeters, the cost-benefit analysts, the systems analysts, in the Defense Department and elsewhere. Whether the parentage turns out to have been legitimate will depend, as is usually the case, on how the offspring turns out.

Higher education may be regarded as a system. It consumes resources, such as money and manpower, and it produces a number of outputs, such as educated people, occupational specialists, professional meetings, and research. There is a production function involved and presumably it is possible to combine the inputs in different ways to achieve desired combinations of outputs, although that has not yet been demonstrated. It is, in this gross sense, like the General Motors Corporation or the Bronx Zoo, unfortunately more like the latter than the former. What makes the system difficult to deal with, among other things, is that its output is not easily specifiable and where we can agree on what the output is it is extremely difficult to quantify it. The inputs are easier to handle, although there are some tricky costing problems, and there is the further complication that it is not always clear whether some parts of the system are inputs or outputs. There is also the problem of which parts represent investment and which parts consumption.

As part of its framework for surveying knowledge about higher education the Public Policy Research Organization has identified six major outputs¹ that are intended to cover, for most practical purposes, the bulk of the output of higher education. They are:

1. *Classification of Youths.* This refers to the much-discussed sorting process carried out by higher education which determines who shall and who shall not be permitted to enter higher education; of those who enter, what institutions they may attend; having begun attendance, who may continue, who may move to a higher ranking institution, who must leave or move to a lower ranking institution; having graduated, who may

¹ This categorization owes much to Robert Bickner and Robert Dubin, both of PPRO.

attempt to obtain an advanced degree and who may not; having made the attempt, who shall and who shall not receive the advanced degree. This sorting process is a major determinant of social structure; higher education tries to carry it out by objective methods, by and large; but of course subjective judgment pervades the whole process and there is no denying that youths from favored families receive special consideration.

2. *Occupational Training.* Most of undergraduate education and essentially all of graduate education is occupational training. Even a purely cultural survey course has an element of occupational training if it requires homework, reports, and examinations. A prime attraction of the bachelor's degree to employers is the evidence it affords that the holder normally disciplines himself to carry out tasks somewhat conscientiously and on time no matter how irrelevant they may seem to be. That's why U.S. Steel does not mind hiring someone who majored in Finnish folklore.

3. *Research.* This refers to generation of new knowledge (ideally) and not to large development programs such as most of those funded by DOD, AEC, and NASA. Departmental research is part of this output as is much of the research funded by private foundations, NSF, and NIH. Research activities belong in this category if society holds the university primarily accountable for their quality; if a contracting agency is held primarily accountable, then the activity is a service and belongs in category 6 below.

4. *Organization of Knowledge.* Higher education devotes a substantial portion of its resources each year to reorganizing the total body of knowledge; it is done mainly by the faculty but students contribute too. The reorganization is required by research findings, by new insights into the interrelations and structures of knowledge, and by society's changing value system which in turn changes emphasis on different segments of the body of knowledge. The reorganization takes place when faculty members formulate their courses each year, when new textbooks are written and old ones revised, when meetings of professional societies take place, when curricular material is debated in the classroom, when journal articles are written that do not deal primarily with new research results, when courses are added and dropped, when students change their pattern of demand for courses.

5. *General Education.* This is the output that you read about in the first few pages of almost any college catalogue but then it usually falls by the wayside as you move into the sorting process and the occupational training. Some four-year colleges are seriously concerned about a general

education; that is, providing each student with a reasonably sound model of how the world operates. Every person has such a model which to him represents the real world and which forms the basis of his decisions throughout life. If it is a realistic representation, he will be a person of good judgment; if not, he will be a person of poor judgment. A college student usually receives very little help from the faculty with this extremely important matter except as bits and pieces of his model happen to coincide with the conventional academic disciplines. Mostly it is constructed from his own experiences and amateur inferences; their integration into a unified structure is left entirely up to him. Except for those bits of specialized curricular knowledge he is not much better off at most colleges than those who do not go to college; he is a little better off because students get help from each other in relating formal course work to their models.

Another aspect of general education which is particularly important to students today has to do with social skills, interpersonal relations, social responsibility, etc. Here also most colleges have little to offer but students do help themselves to some degree.

In sum, there is a noticeable output of general education provided by a few colleges, by a few courses at many institutions, and primarily by the students themselves.

6. *Services.* These are provided mainly by the larger institutions. The institutions carry out engineering and evaluation of weapons systems, space systems, transportation systems and the like for federal agencies. They operate agricultural experiment stations to serve farmers and the agricultural industries; engineering experiment stations serve the manufacturing and construction industries. Tests of drugs and food additives serve the pharmaceutical and food industries and the population at large. Business research institutes serve consumer-oriented industries and small business by maintaining useful statistical indexes and carrying out various kinds of surveys. Local and state governments are served by urban planning studies, water supply studies, sanitary engineering studies, air pollution studies, tax and equalization studies, etc. Institutions are either reimbursed by the agency receiving the service or they receive appropriated funds for performing them as a result of influence exerted by those benefiting from the service. Thus there is no real allocation of resources problem here; the institutions are mainly fulfilling demands of the economy at the market price. Many administrators doubt that the overhead received on government contracts, for example, actually covers the

full overhead; one noticeable, unreimbursable item may be the faculty time spent in acquiring contracts. Still the services cost the institution relatively little if anything and they receive the considerable benefit of keeping in touch with a variety of current practical problems.

This global characterization of outputs is far beyond anything systems analysts are presently trying to do in coming to grips with the resource allocation process in higher education. It is global because the survey is intended not only to discover what is known but to point up important areas where little or nothing is known and substantial research needs to be done. We solicit criticism of this set of outputs.

In a number of major institutions throughout the United States and in Canada, skilled practitioners of the art of program budgeting are attempting to adapt it to university procedures. As was the experience in the Defense Department and later in civilian agencies of the federal and local governments, this is not turning out to be an easy task. While many institutions will offer assurances that they are of course using program budgets, it is our view that no program budget really yet exists in anything but highly partial and elementary form in an institution of higher learning.

One of the healthy by-products of these attempts has been the realization that the kind of information required simply does not exist. In consequence, hand in hand with the attempt to install program budgets has gone the establishment of one or another form of management information system so that the data required to implement the program budget would be made available. It is really incredible how little institutions of higher learning knew about themselves as recently as a decade ago or even today. The computer, of course, has made great progress possible in the collection and storage of simple institutional data. The bulk of the efforts directed toward the installation of program budgets that has been sponsored by the Ford Foundation is now about eighteen months old. In all these cases results are hoped for by the end of three years and in every case the institutions are under injunction from the Foundation to make public their successes and failures, the feasibilities and infeasibilities, and the costs of using the new techniques. Almost without exception the institutions now experimenting with program budgeting are large universities; the comprehensive program budget really seems unnecessary in the smaller institutions, say those with annual budgets of \$10 million or less.

All institutions that we know of that are trying

to reorient their budget structures as indicated above are having great difficulty with the statement of the institutional objectives. Lacking an operational concept of institutional objectives, it is of course conceptually not possible to compare alternative ways of proceeding and to choose a preferred way among them. Does this mean that all is lost in terms of improving resource allocation decisions? Some have urged that it does and that these fancy techniques are, therefore, simply not adaptable to an institution as peculiar and as nonquantifiable as a college or university.

But we argue that all is not lost for the analyst. System-wide objectives are always difficult to specify. Even in a profit-making organization there are questions of the extent of public service that turns up profitable opportunities, differences between long-run potential for profits and short-run profits, and so on, which make the profit criterion far less simple than it appears at first blush. In the nonprofit sector it is even more difficult, but one can argue as Charles Hitch, probably the father of all this, did many years ago in another connection that finding optimal solutions which will move us toward low-level objectives is important, and much easier than seeking optimal solutions that will move us toward unspecifiable high-level objectives. And we think there are a growing number of cases where analysis has resulted in improved decision making in institutions of higher learning, even though specification of overall objectives is not possible. We would like to mention a couple of these briefly.

A year and a half ago Princeton University's trustees decided to begin admitting women as undergraduates in September, 1969, and by now these creatures have become an accepted and appreciated feature of that venerable campus. Behind that decision lies a quite remarkable piece of analysis, published by the University and known, after its principal author, as the Patterson report. No brief description can do justice to this report which may be unique as a faculty analysis of a major issue of educational choice.

In the final analysis there could be only a subjective relation of the educational consequences of the move to the trustees' views of the institution's objectives, since these are not specifiable in operational terms. But the analysis was nonetheless essential in the decision, since it pulled together all the known facts, plus many not known until uncovered by the analysts; thus reducing to a minimum the area to which subjective judgment had to be applied.

To economists, and systems analysts, the most interesting part of the report is Chapter 4, which in a sense is Princeton's first program budget. It

costed the admission of 1,000 women on certain carefully stated assumptions as to what would be permitted to happen to class size and faculty loads. Costing was of a simulated enrollment in actual courses based on courses women elect in other and comparable institutions. The unexpected and remarkable conclusion emerging from this exercise was that 1,000 women would generate an excess of annual expense over income of between \$215,000 and \$380,000, a modest figure in relation to Princeton's annual budget. There would be capital requirements as well, of course, but these carefully derived cost data made it clear that resource requirements would not stand in the way of a decision to "go."

Another interesting example of this sort is the problem that faced the University of California's system when it became clear that its medical school facilities could not meet the demands of the students in the state of California and could not satisfy California's share of the need of the nation for more doctors. One could argue that the addition of more medically-trained men and women is an incorrect objective and that what the country and the state of California need is more paramedical personnel and some institutional arrangements that will push existing medical personnel out of Beverly Hills and into the areas where medical facilities are lacking. But once the decision is made to expand the output of medically-trained people, then a valid question is how this objective may best be met. One obvious answer is to expand the size of the existing medical schools in the university system. The second and competing alternative is to establish new medical schools on campuses not now having them or in independent areas where no University of California campus exists. Within this broad alternative, there are questions of the appropriate size of the new medical schools and the choice of their locations.

Mark Blumberg, of the University's Office of Health Planning, is carrying out a substantial investigation of these questions. We have had the opportunity to see a cost analysis which has just been completed by Paul Wing and Blumberg. It concludes among other things that (1) the annual program costs for undergraduate medical students, interns, residents, and clinical fellows is about \$4,000 each; (2) for postdoctoral science students it is about \$7,000; (3) about 40 percent of expenditures at medical schools is for education and about 60 percent is for research; (4) research-oriented medical schools tend to produce research-oriented doctors; (5) economies of scale are quite small if they exist at all so far as the range of sizes of U.S. medical schools is concerned.

To these findings we add a couple of propositions of our own which we believe can be easily documented: (1) medical schools attract doctors because of the opportunities they offer for internships and residencies; for this and a number of other reasons they are much sought after by localities and the benefits must be real because many localities will pay a sizable price to land a medical school; (2) there is no problem of attracting students; essentially all medical schools have far more qualified applicants than they can accept. If a primary aim of establishing new medical schools is to alleviate the shortage of practicing doctors, then these various considerations might suggest that relatively small medical schools be established; that they concentrate on teaching and do only a minimal amount of research; that they be located in places where there is not already a university campus or other institution of higher education (which would already have provided some of the benefits that would be conferred on the locality by the proposed medical school) and that they be located in places which have great difficulty attracting doctors. This suggested solution is ours, not Mark Blumberg's. We can afford to jump to a hasty conclusion because it is not our problem. The solution obviously puts considerable weight on social benefits; the volume of dollar benefits might be largest if the schools were located in affluent suburbs. The solution may appear to ignore the obvious benefit of locating medical schools at existing universities which arises from joint use of existing facilities such as biology, chemistry, and psychology departments. This benefit is more apparent than real, because medical schools tend to set up their own duplicate departments. Even if that did not occur, we believe that this benefit plus the marginal benefit to the community of adding a medical school is smaller than the benefit conferred on a community which does not have an institution of higher education.

Finally, a perhaps obvious point may be made. When we talk about resource utilization, all resources are swamped by the faculty input. To get materially improved resource utilization one naturally first seeks ways of getting better utilization of the faculty resource. In most institutions today, the faculty-student ratio has reached its present level largely as a result of the happenstance of a number of almost casual factors in the past. The personality of the departmental chairman, bargains struck at a particular time with one or more outstanding faculty member, the condition of the market at a given moment—all of these plus neglect of the matter have brought

faculty-student ratios to where they are.

Minor changes in these ratios have major impacts on institutional budgets. One of us, a couple of years ago, illustrated with an almost naive simulation model, that a small liberal arts college could buy a decade's worth of solvency by shifting its faculty-student ratio gradually over a five-year period from where it was to a point 20 percent lower than it was (some other changes were also a part of the model). The assumption made here was that such a change in the ratio could be achieved without reducing the quality of output. In the case considered this may have been a valid assumption, but in general we know all too little about small and large classes, lectures and seminars, personal relations between faculty and student. Even if we find that results measured by achievement tests do not seem to change with larger classes, this is not really enough.

There are other probable benefits that come from small classes. Two important ones are (1) development of social-intellectual skills and confidence in the informal give-and-take between a professor and his students over intellectual issues and (2) integration of the fragmented information students receive from their narrow discipline-oriented courses and from their experience into a rational model of the world. This second one is probably an extremely valuable benefit; so many student inquiries in the relatively free atmosphere of a small class have to do with exploring the relation of the classroom material to their own view of the world; these explorations with an able adult who has given long and serious thought to that relation are much more important to most students than clarification of the classroom material; they can get the latter from the textbook or from the better students outside of classroom hours.

No instruments are available for measuring these benefits of small classes, as we have indicated. The only real information we have at the moment is the fact that many of the best colleges think small classes are worth spending a lot of money on; budget data from those colleges might give us some kind of crude estimate of the dollar value of small classes as priced in the marketplace. Perhaps "best" in the above sentence might be defined in terms of proportion of applicants rejected. Statistics from, say, the twenty colleges having the highest such proportions might be compared with national averages over all colleges.

We close by commenting on the skepticism that is sometimes expressed about the extent to which knowledge of better ways of doing things can in fact be implemented in institutions of higher education. The point is made that after all col-

leges and universities are essentially run by their faculties, that almost all faculties have great distrust of "efficiency," and that therefore they simply refuse to adopt new ways of doing things if they threaten to disturb at all the old convenient ways of operating. There is of course validity to this line of argument and occasionally faculties have shown themselves unwilling to move even in the face of urgent need. This situation is exacerbated by a growing tendency of faculties to become more oriented toward their disciplines than toward their institutions. Faculty mobility means that if the situation gets unpleasant at one institution, faculty members simply pick up and move to another, taking their research grants with them. This opportunity has doubtless weakened the motivation of some faculty members to be seriously concerned about the health of the particular institution at which they happen to be located at any given moment.

On the other hand, we cannot blame faculties for not rushing to embrace recommendations resulting from simplistic partial analyses of highly complicated issues—analyses which omit factors that are obviously far more crucial than many of the factors included. Sound, thorough analyses will move the great majority of faculty members. The interactions between faculty, students, and administration with respect to proposed devices for improving educational effectiveness will usually bring more and more validity to the evaluation of those devices.

There is a certain irony in the fact that in many cases college and university administrators have been helped by the recent activity of students. In some cases the students have embraced causes which need implementation and have forced them to occur. In other cases, faculties have come to realize that they are simply unable, due to their unwieldy organization, to handle matters which require crisis reaction. Faculties are accustomed to considering matters over periods of months or years and of reaching ultimately a decision to move marginally in one direction or another. The students lack the patience for this sort of management and apparently they have the power to force changes. A number of administrations have shown great skill in using the student unrest in constructive ways in the management of their institutions.

A final word may be in order on a subject rarely discussed, but which is relevant in our view. Although no doubt it may have its advantages, the institution of faculty tenure is a market imperfection that clearly stands in the way of many improvements that many institutions could make.

Born in the depression under conditions that have not existed for decades, the institution of tenure must appear to an objective outsider as primarily a protection of mediocre faculty members. The good institutions behave toward their faculty the way they do, not because of the existence of tenure, but because they know that if they did not behave intelligently they would not be able to attract and hold good faculty people. Able young faculty members know very well that tenure does nothing for them since they are going to be in demand so long as they are good. It is interesting to speculate about what is likely to happen when student bodies realize what the institution of tenure is doing to them. Why this has not yet happened on a substantial scale is not at all clear but it seems almost inconceivable that students will accept the statement from administrators, "We can't do anything about him, since he has tenure." We expect that this ancient institution will soon come under severe attack. The situation is summed up by the young colleague of one of us

who, when offered tenure by his university, answered, "Do I have the option of taking an additional \$500 in salary?" Our guess is that in a bargaining session he would have settled for \$100.

By way of summary and conclusion, we have argued that this complex institution of higher education can and must be analyzed as a system, like other systems. It has its peculiar difficulties and indeed it may be the most difficult of all institutions to deal with in terms of formal systematic analysis. But progress is being made in a number of places and there is little doubt that the supporters of the system are going to demand that much more progress be made as the required resources devoted to the system become greater and greater with the passage of time. It is gratifying that economists, operations analysts, systems analysts, and behavioral scientists are beginning to apply the tools of their trade to this difficult new area. All of us will watch the results with great interest.

DISCUSSION

WILLIAM G. BOWEN: My comments are directed mainly to the paper by Radner and Miller, although I shall also say just a few things about the papers by Hansen and by Kershaw and Mood.

I am particularly interested in the data Radner and Miller have assembled on variations in faculty-student ratios among groups of institutions and over time. However, before commenting on that aspect of their paper, I want to ask one question concerning their study of the demand by students (and their parents) for places at various kinds of institutions: how is student aid treated in their analysis? I cannot tell from the paper itself how this variable has been handled, and yet it is clearly of great importance in many cases.

To return now to the faculty-student ratios, the intriguing question is, of course, what accounts for the observed patterns. The difficulty that Radner and Miller have had to date in finding systematic relationships between faculty-student ratios and other variables, or even in presenting much in the way of conjectures, attests to the complexity of the problem. It also serves as a challenge to the rest of us to come up with some ideas of our own.

In this spirit, I suggest that variations in faculty-student ratios can be understood only in the context of more complete behavioral models for various groups of institutions (the groupings defined more narrowly than in the Radner-Miller paper), with these models designed explicitly to take into account: (1) the educational objectives of the various groups of institutions; (2) the implications for these educational objectives of variations in faculty size and in enrollment, considered separately as well as together; and (3) the financial consequences for the institutions of changes in both faculty size and enrollment. Thinking about the problem in this way leads me to conclude that we are going to need quite different models for different kinds of institutions, and that the model appropriate to a particular institution must reflect the political milieu in which it operates (especially how it gets its money) as well as its educational objectives and current resource constraints. Pursuing this line of reasoning also leads me to conclude that in explaining faculty-student ratios there is no substitute for studying the forces that affect the absolute sizes of both faculties and student bodies, as well as the relationships between these two elements of academic life.

The lack of time at my disposal this morning is a welcome excuse for my own inability to specify in any detail the kinds of models I am advocating. Still, I hope I can at least illustrate the nature of the approach. In the case of the group of institutions I know best (the major private universities), the tendency for the faculty-student ratio to rise somewhat in recent years should be analyzed in terms of three primary considerations. First, as some fields of knowledge splinter and as new specialties develop, a certain number of new faculty positions must be authorized if the institution is to maintain a "critical mass" of faculty in those fields in which it believes it has a commitment to scholarship

and quality graduate education. These new positions are necessary whether or not there is much, if any, increase in enrollment.

Of course, a larger faculty should make it possible to accommodate more students. However, the extent to which increases in enrollment also occur depends on the other two considerations: second, the perceived effects of higher enrollment on the educational role played by the institution within our overall system of higher education; and, third, the consequences of more students for the operating and capital budgets of the institution.

Without taking the time to explain myself, let me simply assert that in the last decade or so it is the financial consideration that has been the more dominant one for most of the institutions in the group I am discussing. The financial effects of higher enrollment will vary according to such factors as: (1) the extent to which additional students can be expected to go into fields where additional appointments are being made; (2) the availability of enough student aid to take care of those students who cannot (or will not) come otherwise; and (3) the extent to which the present physical plant and administrative capacity of the institution must be augmented if enrollment rises.

For this set of institutions, my guess (buttressed by some evidence) is that higher enrollment, even when it accompanies some increase in faculty size that would have occurred anyway, generally means increases in costs, especially capital costs, that are considerably greater than offsetting increases in income. And, unless the missing dollars can be found, an increase in enrollment proportionate to the increase in faculty cannot occur. This will be the case even though such an increase in enrollment might be clearly desirable from the standpoint of effective use of the total educational resources of the nation. In this situation, the marginal costs of the additional students will be far lower than average costs at the institution—and they may also be lower than the costs of providing additional student places at other educational institutions—but none of this will matter if the marginal costs to the institution are still prohibitively higher than the marginal revenue to be derived from the additional students.

There are policy issues of obvious importance here, and they are not unrelated to broader questions of financing higher education that grow out of Lee Hansen's good paper. The Wisconsin-California comparisons contained in Hansen's paper also emphasize the desirability of extending the work that he and Burt Weisbrod have done to other states with different tax structures and different systems of higher education. Also, it seems to me very desirable to include the private sector of higher education in subsequent work of this kind. There are subsidies of various kinds there, too, as a consequence of the use of income from past endowments, as a consequence of student aid policies pursued by the institutions themselves, and as a consequence of the incentives for support of these institutions provided by the tax system.

Let me turn now, in the few minutes that I have left,

to the paper on resource allocation by Kershaw and Mood. In my judgment they have performed a useful service by roughing out a set of six major "outputs" of higher education. I suspect that not all of us will agree with the way each of these outputs is described, but that is partly a matter of differences in taste as well as in perceptions. For my own part, I think that Kershaw and Mood play down the "general education" output too much, and I also think that they overstress the occupational benefits per se of undergraduate education. And, in their discussion of the "services" output, I think that Kershaw and Mood are wrong in asserting that institutions are fully reimbursed for activities of this kind and that, therefore, there is no resource allocation question here. Substantial commitments of their own resources are being made by many institutions as they provide community services, especially in ghetto areas and for the benefit of minority groups.

In the latter half of their paper Kershaw and Mood make a strong case for greater efforts by educational institutions to relate expenditures to program objectives and to make explicit choices among alternative ways of doing things. On the basis of even limited experience with this approach (applied especially to questions of coeducation and to student aid at graduate and undergraduate levels), I endorse this plea strongly. No one should be allowed to believe that greater use of systematic techniques of analysis is going to render judgment unnecessary or depersonalize the decision-making process. On the contrary, rigorous analysis permits a "cleaner" view of the real alternatives and thus permits more people to participate intelligently in the making of policy decisions. In principle this would surely seem to be a good thing—and, at least on days when I can manage a tolerably optimistic view of the human condition, I think that it is. I confess, however, that there are other days when I can make a good case for obscurantism (at least to myself).

Staying with the more pessimistic mood for a moment more, I feel an obligation to challenge what Kershaw and Mood have to say about academic tenure. I challenge them not on their grounds, for tenure surely is a "market imperfection," but on the grounds that it is an imperfection that may be well worth accepting as the price of an important protection against threats to conscience and freedom of expression—threats which, unhappily, seem to me to be on the rise.

ALLAN M. CARTER: As one who has labored in the economics of education vineyard for nearly a decade, variously as grape picker, vintner, and imbiber, I am heartened by the three papers given here. They are not quite evidence of a great vintage year, but they suggest that we are beginning to ask the right questions. "Right" in this context means "relevant," for as a professor-turned-administrator—or "abstractionist" turned "pragmatist"—one sees how critically necessary it is to be able to address the attention of the academic community to structural and procedural questions in the context of informed judgment rather than in the mists of emotion of the new or old politics.

Lee Hansen, in his paper—and in the California study done with Burt Weisbrod—is challenging one old shib-

boleth. The pricing system of higher education, under a dual system of public and private colleges and universities, has made little sense historically; it is even less well adapted to a world of universal higher education. The zero price, or heavily subsidized tuition level, of most public systems was an admirable egalitarian gesture when numbers were small or admissions based on meritocratic principles—but the price of egalitarian admissions policies must ultimately be, I believe, a pricing system based more on ability to pay. The failure to revise public policy along such lines, I am convinced, will gradually destroy the private sector and ultimately result in too few resources being allocated to higher education—and those ineffectively. I find the erosion of the private sector an undesirable path less because of its "privateness" than because it adds a quality of diversity. This is not to make a virtue out of a necessity, but to decry the tendency towards mass-conformity evident in many public systems of higher education. A colleague of mine has a recurring nightmare in which every state has twenty-five universities stamped out of the same standard budget formula, education is compulsory through the baccalaureate, and a merged N.E.A.-U.F.C.T. has imposed universal civil service regulations.

Although Lee Hansen may not recognize himself as the Don Quixote tilting at this windmill, I think we will not make much headway in rearranging higher education in a more rational manner until we have broken out of the old pricing traditions left over from an earlier age. Until the last year the C.U.N.Y. has probably been the classic case of unintended redistribution in favor of middle-income groups, resulting from the mystique of free tuition—it would make a good (unflattering?) comparison with California and Wisconsin. I am glad that Professor Hansen rides forth from a state university lair and cannot therefore be dismissed as a "spoiler" from a well-endowed private institution.

Joe Kershaw and Alex Mood have focused on internal management, and I think we should accept their challenge and invite them to give the finished version of their paper at next year's meetings. Certainly they are dealing with microeconomic problems that cry out for clear analysis, and it is to be hoped that the series of Ford supported projects will help us all to become better householders.

Through no fault of theirs, however, the paper leaves me with a slight sense of unease—as though we were debating in 1969 whether the United States should send a few thousand military advisers to Vietnam. As economists we would probably all share in the authors' wish that universities should reach decisions by rational processes, that we could improve resource allocation by program budgeting techniques, and that management information systems would provide the basis of more sophisticated administration.

I would not dispute these objectives, but I would enter a strong caveat that we are ten years too late if we are merely making the managers more effective. In the academic world of the 1970's an improved technocracy will not be sufficient. Yes, we must broaden the rationality of the system, but we must do so through a

greater community of interests—through much broader representation in the process of decision making.

The authors' use the analogy of the Bronx Zoo. I fear that at times we look more like Noah's Ark with 24-hour visitation privileges, riding out a perpetual storm. If anything, I suspect that institutional objectives are becoming less rational today, and that we are being asked to perform social tasks that colleges and universities are ill-designed to serve. The "outputs" that the Public Policy Research Organization has defined are part and parcel of what many students and faculties are rebelling against. The open-admissions, no-grading, and unrestricted-curricula battles are precisely aimed at preventing the sorting process. The established view has been that education is a differentiating experience; today's radical insists that it become an equalizing experience. Sorting, with its historical connotations of perpetuating the differences of birth and childhood environment, is itself an embattled premise.

Research and services are equally touchy subjects on most campuses, and few faculties can achieve consensus on the balance of general education and vocational training. I believe that universities are infinitely more complex than the Defense Department, and that agreement on institutional objectives is only likely at such a high level of abstraction as to be of limited use. The improvement of management information systems as applied to the university, however, is a most worthwhile effort, provided that the end product is understandable by the broader academic constituency, and will enable it to participate in rational choices.

By all means we should strengthen the administration of our educational institutions, but not just by strengthening administrators. Faculties are too tempted to settle for higher salaries, lower teaching loads, smaller classes and better support services without facing the hard allocation problems, or rationally assessing trade-offs in selecting among alternatives. They must be drawn into that process if we are to avoid the polarization into labor (i.e., teachers) and management. Kershaw and Mood note the trend toward discipline loyalty rather than institutional loyalty which has been so marked over the last twenty years, spurred by academic labor market conditions. The market is now reversing itself dramatically, however, and the academic man is likely to become much less of a twentieth century nomad in the new buyers market. This could lead to a more clearly defined adversary position of faculty and administration—or it could strengthen the partnership. The key may lie in the manner in which we use the improved management information systems and improved budgetary techniques.

KARL A. FOX: I will base my remarks on the paper by Kershaw and Mood and suggest some extensions of their analysis. As space is limited, I will make a number of statements without supporting argument and gloss over some not inconsiderable difficulties.

I am very much in favor of the systems approach adopted by the authors. If anything, I am even more optimistic than they are concerning the speed with which the remaining conceptual difficulties can be

cleared up at the national level and at the level of individual universities.

Conceptual Problems Relating to the Higher Education System as a Whole. On the surface, some outputs of higher education appear to have "economic" values and some do not; some inputs into higher education are included in the national income accounts and some are not. To evaluate all inputs and outputs (economic and noneconomic) of the higher education system in a consistent manner, we need a measure which I will call the "gross social product" (GSP), which has the GNP embedded in it. I would approach this measure through a generalization of consumption theory.

De facto, an individual allocates his time exhaustively among a limited number of subsystems of the society: the family; the economy; the political system; the religious or ecclesiastical system; and the scientific system. Each system has its distinctive medium of exchange.

I assume that an individual has curiosity about his material and social environment and enjoys developing his capacity to exert influence upon it and to make choices concerning his allocation of time among different sectors of it.

I assume that, by the time he enters college, an individual has chosen a general life-style which involves a planned (approximate) allocation of his time among different subsystems of the society and an expected (approximate) level of rewards from each subsystem. Within a broad life-style, he subsequently makes more specific occupational and other choices.

Within each subsystem, the individual allocates his time exhaustively among "behavior settings"; in each behavior setting, he performs a role at a certain level of quality, and higher quality levels typically require larger inputs of his time. The total value of the rewards he experiences from the social system is a function of the rewards he experiences in each behavior setting. A rational individual may be viewed as allocating his time and other resources among behavior settings in such a way as to maximize his expected total rewards. His inputs into the society are his contributions; in equilibrium, his marginal contribution should (as he perceives it) equal his expected marginal reward. By the same kind of accounting convention that sets gross national income (GNI) equal to GNP, we can require for the society as a whole that total rewards equal total contributions: gross social output (GSP) equals gross social input (GSI). We get out of the social system what we put into it. The logical unit of account for the GSP would be dollars, and the ability of the individual to allocate his time and other resources among subsystems should provide the key to a matrix of exchange rates between time, money, and the media of exchange used in the family, political, ecclesiastical, and scientific subsystems.

The higher education system consists of a set of behavior settings. (A classroom containing a teacher and some students would be a behavior setting; a classroom standing vacant would not be a behavior setting in the present sense, because no human behavior was occurring in it.) The outputs of these behavior settings consist of the rewards experienced by all the participants in them.

The gross social product of all behavior settings controlled by the higher education system would be the contribution of higher education to the GSP.

Individuals use university behavior settings differently depending upon their prospective life-styles, which imply self-judgments concerning their abilities to achieve given levels of rewards from different sub-systems of society with given expenditures of time and other resources.

We have some national "market" data concerning outputs of higher education and could no doubt develop more. For example, we have Emanuel Melichar's study of salaries of scientists based on the 1966 National Register of Scientific and Technical Personnel. We could construct output measures for published research. There have been many estimates of the vocational value (career income benefits) of completing an undergraduate college degree. Opportunity cost also gives us a starting point for evaluating the so-called "consumption" aspects of undergraduate and adult (or continuing) education.

Conceptual Problems Regarding Individual Universities and Their Components. Within an individual university, it is clear that various measures of inputs, outputs, and technical coefficients can be developed. It is also clear that optimization techniques can be used in universities (just as in firms), once prices per unit have been assigned to each output and variable input. Even if the prices used in the objective function for an optimization model of a university differed considerably from those which might be appropriate for GSP accounts, they could serve as a basis for internal computation and communication between different levels of administrators in the university. McCamley has demonstrated how the Kornai and Liptak two-level planning approach could be applied to the allocation of "central" resources among the departments of a college if (for purposes of computation) the dean and the department chairmen employed the same set of output prices.¹

Optimization models of a department and its sub-departments or fields could also be used as a basis for discussions among faculty members and perhaps between committees of faculty members, graduate students, and undergraduates majoring in the department. Given an initial set of resource constraints, different sets of relative prices (objective function weights) might lead to different patterns of allocation both of budget resources and of students' time. The computer could not dictate the optimal pattern, but it could demonstrate the sensitivity of optimal solutions to variations in relative prices within ranges that seem reasonable after discussion among the constructively interested parties.

If provisional prices for outputs of higher education were derived at the national level in connection with GSP accounts, they could be multiplied by the corresponding quantity figures to give measures of GSP

produced by individual colleges and universities and by geographical or other aggregates of these institutions.

W. ALLEN WALLIS: It is merely paraphrasing the authors, not criticizing them, to point out that their papers contain no new results that are conclusive. Instead, the papers represent advance accounts of the early stages of research in progress. As such, I found them interesting, stimulating, and important.

The three papers have little in common except higher education. W. Lee Hansen deals with broad and basic issues of public policy in financing higher education. Roy Radner and L. S. Miller deal with certain significant aspects of the sociology of higher education. Joseph A. Kershaw and Alexander M. Mood deal with central problems of managing an institution of higher education.

Hansen finds in Wisconsin, as he and Burton A. Weisbrod found earlier in California, that public financing of higher education "favor[s] by and large the upper-middle and upper-income groups at the expense of the lower-middle and lower-income groups." If this is correct—and I see no reason to doubt it—it is another manifestation of one of Milton Friedman's laws: that public policies purporting to help the poor generally hurt the poor but help those who are not so poor; and it puts higher education on a list that includes social security, minimum wages, labor unions, farm price supports, urban redevelopment, and innumerable other politically popular programs.

Hansen concludes his paper by asking whether "society wants to continue to produce these redistributive effects." While I am not sure what the phrase "society wants" means (indeed, I think it has no meaning), I feel confident in assuring him that public policy will continue to be essentially the same—more because of than in spite of the redistributive effects he reports.

A difficulty in relying on the present method of public financing of higher education, in my opinion, is that less money will be forthcoming than people would be willing to supply if they had a way to provide funds specifically for higher education. State treasuries will be under increasing pressures to keep taxes down and at the same time under increasing pressures to raise spending for all sorts of good causes—for transportation, health, welfare, pollution control, law enforcement, art, music, parks, housing, and so forth. These pressures, it seems to me, are likely to lead to retardation of the growth of public universities. They may lead also, as a measure of economy, to more grants to students for tuition at private institutions. California, Illinois, Wisconsin, and, especially, New York already make such grants. Each student taking advantage of them saves the state several thousand dollars—sometimes as much as \$10,000 for four years. In addition, public financing of higher education by grants to students enables students, their families, and others to add to the grants, thereby channeling directly into higher education money that they would be reluctant to see channeled into higher taxation.

It is not clear from Hansen's paper whether he has taken account of the favorable economic effect on

¹ Francis P. McCamley, "Activity Analysis Models of Educational Institutions" (unpublished Ph.D. dissertation, Iowa State Univ. Lib., 1967).

nongraduates that results from increasing the number of college graduates. I am thinking not of general external benefits, which he does mention, but of the effect on the return to one factor of production when its quantity is reduced relative to the quantity of another factor of production. Increasing the ratio of graduates to nongraduates tends to raise the wages of nongraduates relative to the wages of graduates; thus offsetting a part of the redistributive effect that Hansen finds. He mentions a similar effect from the higher taxes that graduates will pay because of their higher incomes. Still another effect of the same kind arises from gifts to education that are induced by tax exemptions.

The Radner-Miller paper is so preliminary that it is hard to comment on it at this stage. I find myself worried about their data on student-faculty ratios. What do they do about part-time faculty and students? In medical schools, how do they count interns and residents? Since they mention the University of Rochester specifically, I will note that half of our faculty are in the Eastman School of Music and the School of Medicine, where student-faculty ratios are abnormally high—about 5-to-1 in music and 1-to-1 in medicine. Furthermore, Radner's and Miller's study of trends would be confused by places like Rochester, because growth within the period they study has been greater in arts and science, education, business, and engineering, where student-faculty ratios are higher, than in music and medicine. This shifting composition adds an artificial positive component to the trend in the ratio of students to faculty.

Kershaw and Mood take an approach to university management that implies a considerable departure from the traditional view of universities that some of us still cherish. The Kershaw-Mood view will prove, I think, the right one for most institutions in the future. But in the future, with college education becoming as nearly universal as high school education is today, there should be even more diversity among institutions than in the past, when the constituencies and the purposes of all universities were more homogeneous. Thus, some universities should maintain the traditional pattern, even though they will be only a minority.

Traditionally, the university has been regarded simply as an environment in which a great many and a great variety of intellectual activities can take place. The faculty has been thought of as the real heart and substance of the university, not simply a part of its apparatus for serving clients. The faculty, furthermore, has been regarded as essentially a group of independent professional practitioners, more like the partners in a law firm or the doctors who own a private clinic than like the staff of a government bureau—much less like the employees of a corporation.

Under such a view, I do not find it as incredible as do

Kershaw and Mood that institutions of higher learning have little managerial information about themselves, that they have "great difficulty with the statement of their objectives," and more generally that their administrators do not think and act like managers.

Under the traditional view of universities, I find it hard to regard tenure as simply "a market imperfection that clearly stands in the way of many improvements." Tenure may possibly have outlived its usefulness, but if so I think the reasons lie in other directions than those suggested by Kershaw and Mood.

The purpose of tenure, as I see it, is not to confer a privilege on the faculty. Rather, it is something needed by the trustees in order to fulfill their trust. They need to be assured about the faculty—as the authors of the United States Constitution felt they needed to be assured about Supreme Court justices—that they will speak and write the truth as they see it, influenced by no other interest than truth—in particular, not influenced by fear of losing their jobs.

Tenure is under widespread attack today. A valid attack, it seems to me, must show that tenure is not needed to protect academic freedom. That may well be true, but if it is, no one yet has made the case effectively.

To try to make the case, one might first point out that few if any so-called "academic freedom" cases in recent decades have involved beliefs, teachings, or writings in a professor's field of professional competence: they have not involved the functions that tenure is designed to protect. Furthermore, faculty members today have so many interests that are more powerful than mere job security that it is doubtful whether tenure does to an appreciable degree free professors from conflict of interest. Besides, tenure protects against an economic risk that is almost nonexistent in an expansionary period, when the demand for faculty is great relative to the supply. Finally, the arguments for tenure have a validity at a traditional university, where basic research and radical innovative thought are an important institutional purpose, that they do not have at Rattle Snake County Junior College, at a College of Pottery, Printing, and Photography, or at a Jesuit Theological Seminary.

Those are the directions, it seems to me, in which challenges to tenure might be raised.

Despite these remarks, there is no question in my mind that the kind of managerial approach taken by Kershaw and Mood will prove immensely valuable in innumerable ways in all of higher education—at the Harvards as well as the Rattle Snake Colleges.

All three of these papers tackle significant problems of higher education. All indicate fruitful approaches to those problems. I look forward with high expectations to the finished studies.

THE TEACHING OF ECONOMICS

A RADICAL APPROACH TO ECONOMICS: BASIS FOR A NEW CURRICULUM*

By RICHARD C. EDWARDS, ARTHUR MACEWAN, and the
STAFF OF SOCIAL SCIENCES 125
Harvard University

I. Introduction

The purpose of this paper is to outline a radical approach to economics and to suggest how several important social problems might be dealt with in that framework. Our effort to develop a new curriculum is motivated by the conviction that the orthodox approach to economics cannot deal with the important problems of modern society.

Orthodox economic analysis as presented from the elementary course through the graduate seminar is based upon an acceptance of the *status quo* in social relations. Microanalysis presupposes the individualistic ownership and decision-making systems typical of capitalist societies, and in this narrow context the pecuniary behavior of firms and individuals is examined. In macroanalysis, when the aggregate operations of these individual units are the subject matter, attention is focused on the fiscal and monetary adjustments necessary to keep the system smoothly functioning. All in all, the curriculum of modern economics is one of philosophic marginalism: existing social relations are taken as a datum and the problem is one of administering the system by adjustments around the edges.

The marginalist approach is useful only if, accepting the basic institutions of capitalism, one

is primarily concerned with its administration. If one questions the virtue of capitalism as a system, then the basic social relations and the institutions of the system themselves must be subjected to analysis. A new approach is necessary.

The old approach—that which accepts capitalism and is in general the basis of present economics curricula—cannot deal with the problems of modern society. All that the curricula say about the war in Vietnam is how it can be financed more efficiently. The very existence of imperialism is denied. Racism, it is taught, has its origins in personal preferences, and the poverty of blacks and others is “explained” in terms of their low productivity. The destruction of the environment enters the curricula only as an aside when the existence of “externalities” is pointed out as limiting the theory. The subjugation of women, the meaninglessness of work activities, and the alienation of workers are topics which do not enter the curricula at all. Socialist alternatives and the process of revolution are examined only in terms of the value system of a capitalist environment.

It is our contention that such issues—their historical existence, causes, dynamics, and consequences—should be central to a new economics curriculum. This curriculum would reflect the motif of modern American capitalism: conflict and power. Attention would be focused upon the basic economic institutions of capitalism and the class divisions which those institutions foster.

In Sections II and III we lay out a basic substantive argument for a radical approach to economics, which can be summarized as follows. We begin with an analysis of the fundamental capitalist institutions. These institutions function so as to limit the range of social outcomes available; we show how the social problems mentioned above (income inequality, alienation, imperialism, and so forth) are directly attributable to the operation of these institutions. But the basic institutions also confer power differentially, favoring those who already benefit from the

* Social Sciences 125, “The Capitalist Economy: Conflict and Power,” is a course offered in the General Education program of Harvard College. We requested that the course be “cross-listed” in the Department of Economics so that it could count toward concentration credit for economics majors. The Harvard Department of Economics did not grant that request. The staff has met on a weekly basis since the autumn of 1968 to work out the substance of a radical approach to economics and to form the course around that substance. The ideas put forth in this paper are the products of the group. Their expression here has been the responsibility of Richard Edwards and Arthur MacEwan. Other members of the group, which is about half faculty and half teaching fellows, are: Keith Aufhauser, Peter Bohmer, Roger Bohmer, Samuel Bowles, Herbert Gintis, Carl Gotsch, Stephan Michelson, Ralph Pochoda, Patricia Quick, Michael Reich, and Thomas Weisskopf. We are grateful to Janice Weiss for help in editing the paper. This is an abridged version. The full text is available from the authors.

economic system. Therefore, of the limited social outcomes potentially available, there is a tendency to choose those outcomes least conducive to a decent society. We discuss (in Section III) how this power is exercised, particularly through the state, in the service of class interests. In Section IV we conclude with some remarks about methods of teaching. A statement on grading is included as an appendix.

II. *Conflict, Power, and Institutions*

The problems we have cited as providing the motivation for a new economics curriculum—imperialism, inequality, alienation, racism, etc.—directly involve economic conflicts; that is, in each case there are social groups with contradictory economic interests. Conflicts are decided through the exercise of power and through the operations of institutions. More precisely, power—the ability of groups or individuals to resolve conflicts in their favor—is not exercised in a vacuum; rather it is always exercised within a well-defined environment of economic institutions which place strict limitations on its scope of operation.

For example, in the determination of wages in a capitalist society, the institutional environment narrowly confines the scope in which collective bargaining, a process involving power, takes place. First, the bargaining is predicated on the assumption that the struggle is one over distribution of “excess profits”; that is, over what is left after all the “costs” of production (“normal” profits and socially unnecessary expenditures such as advertising, as well as socially necessary costs) have been subtracted from total revenues. In the context of capitalism, the size of these costs is nonnegotiable. Indeed, in the context of capitalism, it would likely be against the interests of the workers involved to cut into these costs because doing so might force the firm out of business. Second, once a negotiated agreement regarding wages has been reached, it can often be vitiated by price increases. Thus, both before and after its operation, power in the bargaining situation is severely constrained by the institutions of the system.

One hypothesis which lies at the core of a radical approach to economics is that basic economic institutions to a large extent determine the nature of social relations and the outcomes of social conflict processes; that is, social decision making is largely organized and effected through the basic economic institutions. This hypothesis involves two questions. First, to what extent do institutions directly determine social relations

and the outcomes of social conflict processes? Second, to what extent is the distribution of power among groups and individuals determined by the structure of institutions?

In considering the role of capitalist institutions, we emphasize as basic (that is, system-defining) institutions the following:¹ the market in labor, in which labor is treated as a commodity and allocated on the basis of the highest bidder; control of the work process by those who own and control capital,² including the concomitant loss of control by the worker over his activities during the hours of work; the legal relations of ownership, by which income distribution is determined through payments to owners for the use of their productive factors; *homo economicus*, the system of personality traits characteristic of and functional to capitalism, including especially the system of individual gain incentives; and the ideology which abstracts and organizes “reality” in such a way as to justify and facilitate the operation of the other institutions.

These institutions create several of the social problems we have mentioned: income inequality, alienation, destruction of the environment, and imperialism. Furthermore, racism and the subjugation of women become functional in a society organized by these institutions. The arguments we shall suggest next are intended to make explicit the links between the operation of capitalist institutions and these problems.

The Consequences of Capitalist Institutions: Income Inequality. Tendencies toward inequality are an integral part of the functioning of capitalist institutions. Consider first the consequences of a market in labor. In order to insure that the vast majority of workers will sell their labor power on the market, it is important that workers not have the option to work for themselves; that is, it is necessary that workers own no factors of production other than their own labor [38, VIII] [49]. As a result, capital ownership must be concentrated in the hands of a relatively small number of nonworkers. Furthermore, as long as material rewards are the main motivation for work, the incentive structure required to induce workers to acquire and apply productive skills must be char-

¹ See Polanyi in Dalton [14], [15] [37] for discussions of the operation of capitalist institutions and their development.

² While “owners” and “controllers” are not necessarily the same persons, as a group they define the goals of the capitalist firm to be profit maximization; on this point see Baran and Sweezy [3] and Solow [52] who counter Galbraith [21]. All we require here is that firms “approximately” maximize profits, or that they do so in the “long run.”

acterized by significant inequalities in labor earnings.³

The capitalist's side of the production process makes similar demands. First, substantial reward differences are needed in order to induce entrepreneurs to perform their social functions as innovators, production organizers, and risk-takers. Second, given economies of scale in production (either technological or those deriving simply from market power, etc.) and given the institutional association between capital ownership and control of the productive process, concentration of ownership necessarily develops.

Finally, profit maximization leads to a rapid rate of technological change. In a market setting, the capacity to exploit profitable innovations depends on the ability to raise the necessary capital. Once inequalities begin to develop (or given historic inequalities), this ability is unequally distributed. As a result, the rewards of technological change, which are often considerable, tend to be distributed to those who are already at the top of the income distribution, lending a further tendency away from equality [50]. The above tendencies of technological progress operate even in the absence of biases which directly reduce the labor share of income in the choice of capital- or labor-saving technology. In a society with more or less competitive factor markets but with collective control by capitalists over research and development, it can easily be shown that the selection of the pattern of technological progress will be to the disadvantage of labor.

These tendencies toward inequality derive directly from the fundamental institutions. We would not argue, however, that the distribution of

income is entirely determined by the operation of capitalist institutions. Income determination is—perhaps above all else—a struggle. One of the principal aspects of this struggle is the effort by groups to increase the prices of the factors they own. Thus, the income struggle can be viewed as a class struggle, where classes are defined in terms of their relation to the means of production.

Alienation. The leisure-labor dichotomy characteristic of neoclassical economic analysis reflects an acceptance of the notion that in general, labor or work activities will be nonfulfilling drudgery undertaken to secure an income, and that creative activities leading to individual development must necessarily be nonwork ("leisure") activities. There is considerable anthropological evidence that this division of life is historically specific to labor-market societies, and that productive activities have not always been separated from creative, developmental ones [14, pp. 19–25].

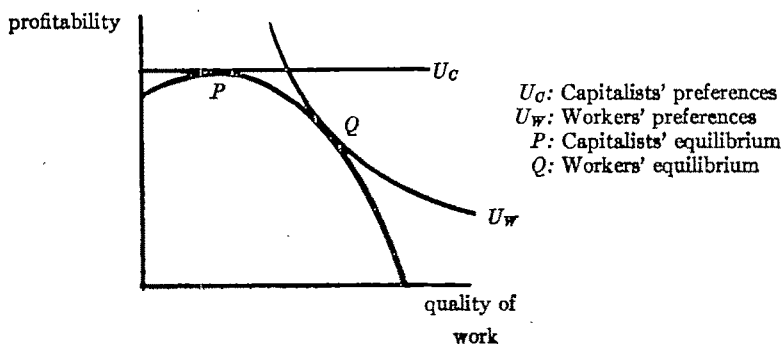
Within capitalist society, the capitalist's control of the work process means that the workers—that is, those who sell their labor on a market—do not determine the technological or social organization of the work process; likewise, they do not determine what product will be produced or what the product will be used for. Thus, the worker is separated or alienated from both his work activities and his product. Likewise, since labor power cannot be separated from the laborer himself, control by the capitalist of the worker's labor carries with it control of the worker's life during the work day.⁴

Under these circumstances, work activities are

⁴ Marx [39] provides the classical statement of the process of alienation. Readings based on the more modern situation include [22] [12] [6] [28] [20].

⁵ More generally, insofar as a trade-off exists between the quality of the work process and maximum profits, in capitalist society the former will always be sacrificed to the latter. This situation is illustrated graphically below.

³ Inequalities in conjunction with personal material incentives exist in some socialist societies as well as in capitalist societies. Meade [41] explains the dual role of prices in a market economy, pointing out that prices which lead to efficient allocation may yield a very undesirable income distribution.



in general neither creative nor self-developmental.⁵ The worker has no intrinsic interest in either his direct activities or their goal, and motivation must then take the form of working for the extrinsic incentive of wages. As pointed out above, for wage incentives to be effective, considerable inequality (and therefore considerable reward for working properly) must exist.

But solving the problem of motivation in an alienated work environment is not left to wages (and the requisite wage inequality) alone. In modern capitalist societies, the United States in particular, the educational system serves the function of preparing workers for the conditions of their employment. The educational system disciplines the work force.

Most people on the job find little use for those cognitive abilities acquired in school, other than the most elementary ones (the "three R's"), or, insofar as they do, they could just as well have learned these abilities on the job. On the other hand, coming on time, following directions and learning to respect authority, learning to work for external incentives (grades), and budgeting time are modes of behavior, affective traits, which the school instills and the job requires. Thus the schools prepare, by experience, their students to function effectively in an alienated environment.⁶

The consequences of alienation are obviously very great. When the organization and purpose of one's major life role—i.e., work—are externally controlled and motivated, that role and life itself tend to become meaningless. Character development and self-expression are distorted and stifled by the work environment and achieve only stunted realization through nonwork activities. Cynicism—towards oneself, towards others, and towards society—tends to be the result.

The worker's alienation can be viewed as a fragmentation of his existence: his working hours are not controlled (arranged, organized, or motivated) by him and are therefore fragmented from

the rest of his existence. His family, his recreation, his intellectual activity are not integrated with work activity which dominates his life.⁷ Furthermore, even his work activities are fragmented: capitalist production drives towards a technology with an ever finer division of labor, so the worker cannot even participate in production of a complete product.

Alienation is not a "cost" to workers which can be recouped through a higher supply price of labor. In the first place, workers are not able to extract higher wages as "compensation" because there are few if any meaningful nonalienating alternatives available to them; that is, alienation is pervasive throughout the capitalist economy. Furthermore, insofar as a variety of work conditions does exist, stratification of labor markets insures that persons in alienating work environments have only similar environments as alternatives (a bureaucrat typically does not have the option of becoming a doctor). Also, the lack of creative, self-developmental work activities insures that workers will not demand such a work environment: their preferences (like everyone else's) are molded by their environment, and the absence of nonalienating alternatives allows workers no basis on which to change their preferences.

Destruction of the Environment. Capitalism is usually credited—by Marx, Schumpeter, and many others—as being a system which attains maximum output expansion from a given resource base. Markets and *homo economicus* prove to be powerful tools for organizing an economy towards growth. The ideology of capitalism, in turn, places high value on the rise of material output.

The acceptance of aggregate output per se as an indicator of welfare, however, would be at best a questionable procedure. The prices at which aggregate output is valued are reflections of the existing distribution of income and the preferences engendered by the system. They therefore cannot be endowed with any objective welfare meaning. Furthermore, the very process of output expansion has consequences outside of the market which are detrimental to social welfare.

The capitalist growth process has historically involved the fracture of community. This process takes many forms, including current phenomena: traditional, nonalienating work processes are destroyed by competition with modern industry; agricultural communities are decimated by the

Of course, in a society where workers' preferences dominated, the transformation locus might look very different, due both to a different motivational orientation of workers and conscious development of new technologies consonant with higher quality work.

⁵ More on this argument, especially with regard to the role of grades, is provided in the appendix. Useful readings concerning the function of schools include [24] [48] [27] [19]. It should be noted that what has been said about alienation and education in the capitalist environment could also be true in other modern, highly bureaucratized and organized societies. Thus, while the elimination of capitalism is a necessary condition for the elimination of this problem, it is not a sufficient condition. On the personality requirements of bureaucracy, see [22] [42].

⁷ On the issue of fragmentation of life, see readings listed in the preceding footnote as well as Polanyi in [14] and [29, especially pp. 243–68].

introduction of new technology; cities grow and decay depending upon the vagaries of the market; urban inhabitants experience anomie because their communities are functionally fragmented. Community is not a good which can be produced for market sale.

Because the capitalist controls the work process and his goal is profit maximization, there will be no tendency to minimize costs which fall on others. Indeed, for any given level of costs, there will be an effort to maximize the share of costs borne by others. These extramarket costs take the form of fracture of the community, water and air pollution, congestion, "urban sprawl," etc.—a general destruction of the environment which cannot be viewed as a secondary issue but one of dominant importance in the society. Furthermore, capitalists' efforts to choose technology and to organize production so as to minimize their own, but not social, costs insures that the importance of the problem increases over time. (The rich are often able to protect themselves from pollution, by zoning for example.) Thus the rise in concern regarding problems of pollution is in no way surprising, nor should it be surprising that anti-pollution groups make headway only when the problems become severe. To halt the destruction of the environment, it would be necessary to restrict seriously the operation of basic capitalist institutions. Thus human needs become subordinated to the needs of the market and to capital expansion.

Imperialism. Subordination of human needs to the needs of capital expansion has been a pervasive characteristic of capitalist growth, and this process has been carried out, not only domestically, but also on an international scale. The geographic spread of capitalism derives from the operations of its basic institutions. First, the individual gain rationale of capitalism leads constantly in search of new sources of profits. For the firm, continued well-being depends upon finding new, profitable uses for its previously accrued profits. Second, the opportunities for assuring the availability of such investment opportunities are greater, the greater is the geographic scope of the system.

The spreading of the system has been a characteristic of capitalism throughout its history. It has involved breaking down the restrictions on the operation of the market, on the capitalist control of the work process, and on the system of individual gain. Earlier, the problem was one of creating nations and then spheres of influence. Today, when one capitalist nation has become dominant, the problem is one of integrating an international capitalist system. This integration

means at a minimum that the nefarious aspects of capitalism—inequality, alienation, destruction of environment—are spread, or, insofar as they already exist, they are maintained. However, because this integration takes place under the dominance of the business interests of an advanced capitalist nation, the output expansion capacity of capitalism is not necessarily transmitted to the poorer countries. First, simply the operation of comparative advantage which operates in an integrated capitalist system would inhibit industrialization and growth in poor countries. Second, the monopolistic conditions of business in the United States allow even less opportunity for development in poor countries. Because of its power, which operates both within and outside the market, U.S. business (or business from other advanced capitalist nations) is able to preempt investment opportunities and inhibit the development of a historically progressive industrial bourgeoisie in the poor countries. Finally, the interests of international capital require the maintenance of a "favorable investment climate" and so the state power of advanced countries is used to prevent radical political and social change in poor countries.⁸

Racism. Racism in the United States can be seen as functionally supportive of the interests of capital. In our discussion of alienation, we asserted that the segmentation and stratification of the labor force is one mechanism by which labor is prevented from obtaining a higher wage for undertaking less desirable jobs. Black people clearly comprise one of the most oppressed segments of the labor force in the United States. They are restricted to the most undesirable jobs and they are paid the lowest wages.

It is often alleged that white workers benefit from racism and that the losers are the capitalists who are prevented by racism from hiring blacks. Such an allegation is true, if at all, only in the static sense when total labor income is fixed. The division of the labor force by race, however, weakens the position of workers as a group, and their share of income is consequently reduced. Demands by white workers are attenuated by the threat of being replaced by workers from the black labor pool.⁹

⁸ Useful references putting forth the radical analysis of imperialism include [56] [36] [18] [2] [32]. On the history of U.S. imperialism, see [31] [61] [47].

⁹ Michael Reich, in his investigation of the relationship between racism and class divisions, has found that racial inequality between blacks and whites (as measured by the difference between white and nonwhite median incomes) is significantly and positively related to class inequality among whites (as measured by the Gini coefficient for white incomes). See [53, Chap. VII and *passim*].

Furthermore, antagonisms of white workers are directed by racism towards blacks rather than towards employers; that is, institutionalized racism dilutes awareness of class divisions. When the animosity of white workers is directed against blacks, the white workers see themselves as having a stake in the system—they are not at the bottom.

The Subjugation of Women. The segmentation of the labor force, of course, involves many divisions other than race. Extensive division by "skill" and education categories is of obvious importance. The division by sex and the concomitant subjugation of women pervade the entire society. In this paper we will only point out that there are many parallels to racism, and the points we have made above, especially those regarding the functional role of stratification in allowing low wages to be paid for undesirable work, are again relevant. Indeed, for most of their labor, namely, housework, women receive no wages as such and have very little choice in the matter [46] [5].

Summary. In this section we have suggested arguments regarding the relationship between important social problems and the institutions of a capitalist society. We believe that the general analysis forms a basic component in a radical approach to economics. The points we would like to emphasize in summary are the following:

1. The negative characteristics which we have ascribed to capitalism are completely compatible with successful, rapid expansion of output. Indeed, it is in the very process of yielding a maximum output (maximum profit) situation that the institutions of capitalism yield income inequality and alienation, for example. Thus, the analysis is only in small part based upon the occurrence of business cycles, unemployment, etc. In essence, the critique is fully applicable when the capitalist economy is in boom. And however one may value output versus other variables which contribute to welfare, it seems obvious that the greater the output, the less valuable it is relative to the other variables. Thus a capitalist society becomes increasingly less tolerable.

2. The core institutions of capitalism interact in such a way as to determine social relations and circumscribe the outcomes of conflict situations. First, these institutions work in a parallel fashion to produce a class society. Reliance on individual gain incentives, capitalist control of the production process, and the legal relations of ownership insure that capitalist development will produce division into classes. Second, the core institutions are highly interdependent in that they sustain and facilitate the operation of each other; the functional interrelations are such that severe

alteration of any one is incompatible with maintaining the others.

3. The acceptance of capitalist institutions carries with it certain constraints on the functioning of the society. These constraints usually take the form of conflicts between alternative social needs. For example, the trade-off between output growth and income equality exists within the context of capitalism. The trade-off is created by the functional requisites of the institutions. But we can easily imagine a society in which such a trade-off would not exist. Such would be the case if men worked because they cared for and felt on equal terms with the entire community rather than if they worked for direct personal gain. Other examples of such generally accepted trade-offs, which are artifacts of capitalism, include income growth versus a meaningful work environment, employment versus stable prices, private versus social costs, public versus private consumption, and income versus leisure.

4. The core capitalist institutions tend to subordinate other institutions to serve their needs. We have illustrated with the case of education (other examples would be the family and religion) how other institutions serve the needs of the economy. In this sense, capitalism may be characterized as an "economic society."

There are, of course, limits to the extent to which capitalist institutions shape society, and a curriculum based only on the argument to this point would certainly be deficient. Thus, we now turn to a consideration of the exercise of power—in particular, the role of the state—in the context of capitalist institutions.

III. *The Exercise of Power*

Class Divisions in Capitalist Society. As we have noted, the development and operation of capitalist institutions divides society into classes. First, class division is a prerequisite for the effective organization of the institutions: most of the population must be reduced to worker status while simultaneously a capitalist elite is created and its existence justified. Second, the basic institutions function so as to augment the wealth, power, and privilege of that elite.

The analysis of economic institutions which leads to these conclusions provides a basis for examining the exercise of power—the ability of groups to resolve the outcomes of social conflict processes in their own favor. First, the analysis provides the working hypothesis that economic organization is the basis of power. Second, the analysis emphasizes that the different classes have conflicting interests with regard to the maintenance of the existing social relations. Together,

these statements would lead us to hypothesize that power in a capitalist society is dominated by the capitalist class, and since social conflict may lead to instability in the institutions themselves, the class exercises power primarily to maintain the institutions which function in its favor. The intervention of power—to deflect political threats, depoliticize class conflict, and so forth—assures the smooth functioning of capitalism.¹⁰

The Operation of Ruling Class Power—The State. An example of the interaction between the operation of institutions and the exercise of power is provided by the recent history of welfare programs. As we have pointed out in Section II, an unequal income distribution results from the functioning of the labor market, the system of individual gain incentives, and the linking of income to ownership and sale of productive factors. There are, however, several secondary forces which exacerbate inequality, and the reality of capitalism is even worse than the model. First, there are many family units which own no salable labor or other factors of production: the sick, the aged, the disabled. Second, there are those who own labor power but who are discriminated against in the labor market: blacks, other non-whites, and women. Third, income inequalities are exacerbated by unequal access to activities through which labor quality is "improved" (e.g., schooling and apprenticeship). Fourth, unemployment is always present in a capitalist system, and its incidence falls heaviest on the groups already at the bottom of the income ladder.¹¹

This situation poses a threat to capitalism.

¹⁰ While we argue that power is dominated by the capitalist class, that is not to say that it monopolizes power or that its rule is unrestricted. Furthermore, capitalists need not monopolize decision-making positions nor must they operate according to an articulated schema in order to be dominant. The existence of an ideology which favors capitalist interests and a sufficiently pervasive common set of objective self-interests among capitalists serves to assure that decisions will be in their favor. It is in this sense that we can identify the capitalist class as a ruling class. The dichotomous division of society into workers and capitalists obviously involves a simplification. Other groups (e.g., highly paid professionals, land-owning farmers, etc.) exist who cannot readily be identified directly with either class. However, we use the term "worker" broadly to identify all who sell their labor power on a market and therefore the class categories extend to most of the population. Furthermore, our preceding analysis of capitalist institutions and our analysis below of the exercise of power lead us to the conclusion that these are the most important groups to study for understanding social change. On the American ruling class, see [54, Chap. 9] [45] [16]; on the nature of classes, see [57] [8] [4].

¹¹ Miller [44] provides a good description of these secondary forces. For a left critique of Miller, see [30].

Those affected have no stake in maintaining the system and become unruly. The preservation of capitalism requires that the misery of poverty be alleviated, or at least that something be done about its appearance. Yet an attack on the basic causes of the problem, the functioning of the basic economic institutions, is ruled out. For example, an adequate welfare program would interfere with work incentives; it would conflict with the principle that income is a payment for productive factors. Therefore, political power is focused on the secondary factors and symptoms, but the basic processes remain unaffected. Old age pension programs are established; equal opportunity employment regulations are legislated; manpower training programs are set up; unemployment compensation schemes are developed. Even if such programs were successful on their own terms, they could eliminate only the most severe aspects of inequality and poverty. In fact, most of these programs fail to achieve their own modest objectives.

Opposition to system-preserving welfare programs derives not only from their conflict with the institutions. Often, interest groups within the capitalist class or powerful professional groups are hurt by welfare legislation. Thus, the A.M.A. battles against medical care; housing developers oppose public housing programs and city planning; the automobile companies work to keep public transit facilities inadequate; textile employers subvert equal employment opportunity legislation. These are cases where class interests and self-interest seem to conflict. While the ruling class as a whole would benefit by establishing an ameliorative program and thereby securing its position, some of its members would be hurt. Thus because ruling class solidarity (see below) is at least as important for the preservation of that system as is preventing disruption by the poor, inadequate welfare programs are the outcome.

Welfare programs are but one example of ruling-class functioning—taking action, compromising within itself, absorbing discontent—carried out through the state. Other revealing examples are public education, tariff policies, financing of research programs, agriculture and transportation subsidies, and the structure of taxation. We believe that these operations of the state are best understood if the state is viewed as basically operating in the interests of the capitalist class.¹²

¹² Sweezy [55, Chap. 13] provides a good statement of this view of the state. The classic argument is provided by Lenin [33, especially Part I] [34].

The Priorities of the State. If, as according to our hypothesis, the state is dominated by the capitalist class, then the operations of the state should reflect the needs of the capitalist class. In modern capitalist states, when the basic institutions have been thoroughly established, the maintenance and preservation of these institutions upon which the structure of class and privilege depends is of the greatest importance to the capitalist class. The uninhibited operation of the economic institutions will continue to bestow power, wealth, and prestige upon the capitalists. They do not need the state to enhance their position, only to assure it.

The system-preserving function of the state is evident in several areas. A continued threat to capitalism has been the failure of the economy autonomously to generate adequate aggregate demand. This failure has brought recurring crises with substantial unemployment. In spite of once seemingly inviolable ideological objections to the contrary, the state has assumed the function of demand regulator. Such regulation does not eliminate unemployment, but simply reduces it to levels which are not system threatening.

A second system-preserving function of the state has been its decisive role in obfuscation and suppression of class conflict. This is accomplished through suppressing system-threatening groups (e.g., the Wobblies, Black Panthers), by deflecting their demands for structural changes into acceptable material demands (e.g., labor union economism, black capitalism), or through ameliorative programs. If we may modify the jargon of public finance, state actions such as suppression or amelioration may be viewed as "class goods." When the challenge posed by workers becomes severe, no single capitalist can protect himself. Were he to give concessions to his workers, his competitive position would be endangered. To employ private armies has been possible but highly inefficient. Thus, action by the capitalists as a class is necessary.

The enormous military establishment provides another example of system-preserving state operations; as such, it performs a dual function. First, it provides the rationale for huge expenditures which serve to maintain aggregate demand without threatening the security or position of any group in the ruling class. For example, social welfare measures often do threaten such groups. Second, as the capitalist system becomes increasingly an international system, the military directly protects the far-flung parts of that system [3].

The response of the state to changes in the

process of production which require more highly developed labor, illustrates a second priority of the state; namely, the creation of new institutions. The rise of mass education in the United States has occurred in response to the need by industry for a skilled work force.¹³ Because workers are not tied to particular employment, individual capitalists cannot invest in the general training of workers and expect to appropriate the returns. Thus, capitalists turn to the state to provide a skilled work force. When education is handled by the state and portrayed as social welfare, it is paid for by general tax revenue rather than by the capitalists themselves [35, Chap. 3] [51] [43] [9] [60].

The structure of the educational system betrays its class-oriented genesis. Mass education in the United States covers a vast quality range, and a positive association has been established between parents' incomes or class and the quality of public education which children receive. If, as seems reasonable, the benefits of education are correlated with the quality of that education, then the class bias of U.S. education is obvious. Thus the educational system operates to reinforce the class bias of the core economic institutions.¹⁴

There is a further aspect of the educational function which reveals its class bias; namely, its role in transmitting ideology. Students are taught a view of society which justifies the *status quo* and which poses efforts for change as unnecessary or futile.

The primacy of the roles of the state in preserving the system and in developing new institutions to meet changing circumstances should not obscure the fact that the state also intervenes di-

¹³ See, for example, [11] [13, pp. 23-57].

¹⁴ The relationship between inequalities in schooling and inequalities in the total income distribution may be expressed as follows. Let Y represent individual income, K represent earnings from capital, and L represent earnings from labor. If we let the variance of Y over the mean Y represent our measure of income inequality, it can easily be seen that

$$\frac{\text{var } Y}{\bar{Y}} = \left(\frac{\bar{K}}{\bar{Y}}\right) \frac{\text{var } K}{\bar{K}} + \left(\frac{\bar{L}}{\bar{Y}}\right) \frac{\text{var } L}{\bar{L}} + \frac{2 \text{ covar } (K, L)}{\bar{Y}}$$

On the plausible assumption that most of the inequalities in labor earnings are due to inequalities in skills, education, and the general socialization process, we see that inequalities in schooling may contribute to income inequality, even where school inequalities are not associated with inequalities in capital ownership. However, note that the last term on the right-hand side of the above expression represents the contribution to total income inequality of the degree to which inequalities in capital earnings are associated with inequalities in labor earnings. Given the social class inequalities of our educational system, we expect the covariance term to be positive.

rectly in the economy to benefit immediate interests of capitalists. The most significant realm—in quantitative terms—where the state intervenes is in military and space spending, which we discussed above.¹⁴

Another example of direct intervention, one which illustrates the case particularly well, is the government's relation to the agricultural sector. The general picture of what has happened in agriculture is well known. Wages in agriculture have remained low and unemployment high. Subsistence farmers have been unable to survive. The rural poor have been forced into the urban ghettos, supplying the low-cost labor force for industrial expansion. All the while, large farmers have received subsidies, price supports, and protection [1].

Furthermore, the very process which creates the agricultural problem is exacerbated by government programs. Government expenditure on agricultural research and extension has played a significant role in raising agricultural productivity at a more rapid rate than general productivity and has thereby contributed to the mass-dislocation of rural workers and subsistence farmers. Those statistical studies which are available confirm casual empiricism: the overall impact of the government in its agricultural programs has been to increase inequality within the agricultural sector [7] [10].

The point is, however, not only that the process has worked toward increasing inequality but that it is the large owners of property—of the agricultural means of production—who benefit. Their benefit is derived directly from the programs which have been developed for "helping agriculture." Payment for unused land is of no help to rural laborers. Price supports for marketable surplus is of no help to subsistence farmers. Government subsidies for capital-augmenting technical change have the same class bias.

Military spending, agricultural subsidies, and other such programs provide ample ammunition for the muckraker. However, in terms of their importance in the overall operations of the state, we believe they are not of highest priority. Their position is behind the system-preserving and secondary-institutions-creating roles of the state. Nonetheless, when studied as a group, these actions of the state which directly enhance the privilege of the capitalist class reveal the basic character of the state in a capitalist society and provide a useful starting point for the analysis of power.¹⁵

¹⁴ For a documentation of the subsidies provided to military contractors, see [59] and Joint Economic Committee (1969).

Cohesion of the Ruling Class. The term "ruling class" may evoke the image of a small, conspiratorial group which coldly calculates the oppression of the poor and its own gain. The actual functioning of the capitalist ruling class in the United States cannot, however, be well understood in such terms.

A class operates as a class in a number of ways. First, the class can be conscious of itself as a group with common objective interests, and can function cohesively on the basis of that consciousness. Second, the class can hold in common a value system or ideology which justifies the class's position and serves as a guide to action. Third, the class can coalesce on specific issues which serve the interests of some of its members if the favor is returned when the special interests of other members are at issue.

In general, it is difficult to distinguish which of these three mechanisms is at work at any given time. In the case of the United States, all three mechanisms operate. For example, elite schools, class-segregated neighborhoods, and social clubs tend to instill in ruling class members a sense of identity and of their separateness from the rest of society. Thus, they become aware of their special stake in the *status quo* social relations and consciously work for the stability of the system. Obviously, if aware of their own position and if working toward a common goal, the members of the ruling class need not "conspire" to assure behavior in their common interest.

On the other hand, the very strong capitalist ideology in the United States tends to make class consciousness per se less important. A set of values that justify the position of the capitalist class, the basic institutions of capitalism, and the *status quo* in general provides a guide to action. Indeed, the prevalence of the capitalist ideology not only assures common action by members of the capitalist class but means that others will cooperate to serve capitalist interests above their own. This is the case, for example, when white workers accept racism and reject a working-class consciousness.

On many issues, logrolling furthers the class interest. This occurs when each group within the capitalist class structures its own policies so that they do not come into conflict with other groups within the class, expecting (and receiving) such cooperation in return.

¹⁵ It would be consistent with this theory if the state were to take some actions which, in terms of their direct impact, increased income equality, provided these actions could be interpreted as serving the stability of the system. Lenin [33] analyzes the ten-hour day legislation in these terms.

These mechanisms which tie a class together should not be confused with the objective identity of the class itself. The capitalist class in the United States is a ruling class. The degree to which it has consciousness, a strong ideology, and internal cooperation determines how successfully it can rule.¹⁷

IV. On Teaching and Practice

So far we have limited our discussion to the content of a radical course. Radical substance, however, is only part of a radical approach to economics. First, college and university teachers work in one of the centers of radical social activity in the United States. Radical teachers should, therefore, view their own work as part of a wider radical movement. They should design their courses to be relevant to the concerns and needs of that movement. Furthermore, radical economists cannot be isolated as academics. Only by taking part in the activities of the radical movement can they integrate their scholarly work with the concerns of the movement.

Second, teaching style is also relevant to the radical approach. In Section II we suggested the argument that one of the functions of the educational system in the United States is to prepare students for the authoritarian and repressive

conditions of the work place. Obviously radicals cannot allow themselves to be part of such a socializing process. It is necessary for our own liberation as well as for the students' welfare to break down authoritarian relations in the classroom. The function of lectures, for example, in which one person talks at a mass of students, is, in general, antithetical to radical teaching. Likewise, radical teachers should challenge the grading system and the role which grades play in providing an external incentive analogous to and preparatory for the wage system (see Appendix).

In other words, teaching with a radical approach to economics is not simply a matter of putting forth a certain interpretation of United States capitalism. A radical approach carries with it certain lessons for change—lessons which must be followed in our teaching of economics.

APPENDIX

*Statement on Grading by the Staff of Social Sciences 125 to the Committee on Educational Policy**

Grades serve a number of functions: they establish a system of incentives for students, they structure the nature of social relations in the educational process, and they provide needed information within and outside the university. We argue that the incentive function of grades and their effect on classroom relations are inimical to learning. Furthermore, we believe that the informational role of grades could be served through alternative mechanisms which would promote rather than hinder learning.

Incentives. Learning should take place for desirable social ends and for the intrinsic enjoyment of learning. The grading process establishes an undesirable reward structure in which obtaining a high grade becomes the motivational force. The indirect reward of a good grade replaces the direct satisfaction from the process of learning or the resulting knowledge as the final objective of many students. Such an incentive structure is undesirable in and of itself.

However, the role of grades in educational institutions cannot be fully understood as long as attention is confined to the universities alone. Grades function to socialize students into the work force. On a job, workers do not obtain satisfaction from an intrinsic interest either in the process of production or in the resulting product of their work. Nor do they obtain satisfaction from the social usefulness of the product. Instead, they are motivated by the prospect of an external reward—wages received in exchange for labor powers. In the workplace, the need to substitute external incentives for intrinsic interest arises because of the separation of

¹⁷ Readers of this paper may well ask to what extent we consider our approach Marxist. The following quotation seems relevant:

"... When asked whether or not we are Marxist, our position is the same as that of a physicist or a biologist when asked if he is a 'Newtonian' or if he is a 'Pasteurian'."

"There are truths so evident, so much a part of people's knowledge, that it is now useless to discuss them. One ought to be 'Marxist' with the same naturalness with which one is 'Newtonian' in Physics, or 'Pasteurian' in biology, considering that if facts determine new concepts, these new concepts will never divest themselves of that portion of truth possessed by the older concepts they have outdated. Such is the case, for example, of Einsteinian relativity or Planck's 'quantum' theory with respect to the discoveries of Newton; they take nothing at all away from the greatness of the learned Englishman. Thanks to Newton, physics was able to advance until it had achieved new concepts in space. The learned Englishman provided the necessary steppingstones for them."

"The advances in social and political science, as in other fields, belong to a long historical process whose links are connecting, adding up, molding and constantly perfecting themselves. . . ."

"The merit of Marx is that he suddenly produces a qualitative change in the history of social thought. He interprets history, understands its dynamic, predicts the future, but in addition to predicting it (which would satisfy his scientific obligation) he expresses a revolutionary concept: the world must not only be interpreted, it must be transformed." The statement is by Ernesto "Che" Guevara [25].

* The petition which follows was submitted to the Committee on Educational Policy of Harvard University by the Staff of Social Sciences 125 during the first term in which the course was offered. The petition was rejected.

the workers from control over the production process and its products. Grades play an important role in preparing young people for this kind of work environment. We object to both an economic system and an educational system which operate in this manner.

The content of the knowledge acquired by a student is also affected by grades. Those aspects of any subject matter which can most easily be reduced to a single one-dimensional measure increase in importance—such as, for example, factual and quantifiable data. In the choice of paper topics, preference is given to small questions which can be easily researched and for which a complete answer can be developed in the limited time available. In order to assure a short-run payoff, the student tends to minimize risk by restricting his field of inquiry. As a result, the larger framework and context of his studies is taken for granted. Thus grades play a significant role in the perpetuation of the *status quo* in social inquiry. To this we object.

Structure of Classroom Social Relations. The power to give grades provides professors with a sanction for the exercise of authority in the educational process. Grades promote acquiescence and conformity among students and exempt teachers from the necessity of being relevant, interesting, and well prepared in their classes. Students refrain from criticizing mediocrity and dullness in part because of the fear of jeopardizing their grades and in part because the process of grading has diverted attention away from learning itself. (We do not raise here the possibility that grades inspire political conformity between students and professors.) In general, the authoritarian relationship between teachers and students in the classroom is inimical to learning, and for this reason too we oppose grades.

Information. The principal external consumers of the information contained in grades are employers and graduate schools, who need to identify the students they most prefer and the ones they least prefer. Grades provide employers and graduate schools with a costless means of ranking students for their own purposes. But education should not be made subservient to their needs, particularly since grades interfere with the learning process. Graduate schools and employers could devise their own mechanisms of evaluation and selection if students were not graded, as already happens with students from a number of colleges, such as Antioch, which do not grade.

Grades are also used to fill informational needs within the university. Students use grades to obtain feedback from their instructors on performance in class. Faculty members use grades from previous courses as guidelines for admitting students to their own courses. The administration uses grades in allocating financial aid. Although we object to this last use of grades, we do feel that information on student performance can be useful both to the student and to the teacher in the educational process. The use of a summary letter grade is simply not the best means for fulfilling such informational needs.

For all these reasons, we find the grading process abhorrent, and we intend to substitute other mechanisms to perform those functions of grades that we feel

should be retained. In our course we will prepare written evaluations of each student's work. The evaluation will be available to the student and to others if the student so requests. Further, we plan to arrange individual meetings between student and instructor during the semester. Finally, the organization of the course into small sections automatically provides continuous feedback to the students.

Therefore, we petition the Committee on Educational Policy to remove the grading requirement from our course. Further, we ask for a public hearing with the C.E.P. concerning both our petition and the general role of grades at Harvard. We would like to raise at that point the arguments for the complete elimination of grades from the Harvard educational process.

REFERENCES

1. SIDNEY BALDWIN, *Poverty and Politics* (North Carolina Press, 1964).
2. PAUL A. BARAN, *The Political Economy of Growth* (Monthly Review Press, 1957).
3. PAUL A. BARAN AND PAUL M. SWEEZY, *Monopoly Capital* (Monthly Review Press, 1966).
4. REINHARD BENDIX AND SEYMOUR MARTIN LIPSET, *Class Status and Power* (The Free Press, 1966).
5. MARGARET BENSTON, "The Political Economy of Women's Liberation," *Monthly Rev.*, Sept., 1969.
6. ROBERT BLAUNER, *Alienation and Freedom* (Univ. of Chicago Press, 1964).
7. JAMES BONNEN, "The Distribution of Benefits from Selected U.S. Farm Programs," *Rural Poverty in the United States* (National Advisory Commission on Rural Poverty, May, 1968).
8. T. B. BOTTOMORE, *Classes in Modern Society* (Random House, 1966).
9. SAMUEL BOWLES, "Towards an Educational Production Function," forthcoming in W. Lee Hansen, ed., *Education, Income, and Human Capital* (N.B.E.R., 1970).
10. DAVID BOYNE, "Changes in the Income Distribution in Agriculture," *J. of Farm Econ.*, Dec., 1965.
11. FRANK TRACY CARLETON, *Economic Influences upon Educational Progress in the United States, 1820-50* (Univ. of Wisconsin Press, 1908).
12. ELY CHINYO, *Automobile Workers and the American Dream* (Beacon Press, 1965).
13. LAWRENCE A. CREMIN, *The Transformation of the School* (Vintage Books, 1964).
14. GEORGE DALTON, ed., *Primitive, Archaic, and Modern Economies: Essays of Karl Polanyi* (Doubleday and Co., 1968).
15. MAURICE DOBB, *Studies in the Development of Capitalism* (International Publishers, 1963).
16. G. WILLIAM DOMHOFF, *Who Rules America?* (Prentice-Hall, Inc., 1967).
17. ROBERT DREBEN, *On What Is Learned in School* (Addison-Wesley Pub. Co., 1968).
18. ANDRE GUNDER FRANK, *Capitalism and Underdevelopment in Latin America* (Monthly Review Press, 1967).
19. EDGAR Z. FRIEDENBERG, *Coming of Age in America* (Random House, 1963).

20. ERICH FROMM, "Individual and Social Origins of Neurosis," *American Soc. Rev.*, June, 1944.
21. JOHN K. GALBRAITH, *The New Industrial State* (Houghton Mifflin Co., 1967).
22. HANS GERTH AND C. W. MILLS, eds., *From Max Weber: Essays in Sociology* (Oxford Univ. Press, 1958).
23. HERBERT M. GINTIS, "Alienation and Power: Towards a Radical Welfare Economics" (unpublished Ph.D. dissertation, Harvard, 1969).
24. ———, "New Working Class and Revolutionary Youth" (forthcoming in *Socialist Revolution*, 1969).
25. PAUL GOODMAN, *Like a Conquered Province* (Random House, 1968).
26. ERNESTO GUEVARA, "Notes for the Study of the Ideology of the Cuban Revolution," *Studies on the Left*, Vol. 1, No. 3, 1960.
27. ALEX INKELES, "Social Structure and the Socialization of Competence," *Harvard Educa. Rev.*, June, 1966.
28. ERIC AND MARY JOSEPHSON, eds., *Man Alone* (Dell Pub. Co., 1962).
29. KENNETH KENISTON, *The Uncommitted* (Dell Pub. Co., 1960).
30. GABRIEL KOLKO, "On Blaming the Poor for Poverty," *New Politics*, Vol. 3, No. 2, 1965.
31. WALTER LAFFEBER, *The New Empire* (Cornell Univ. Press, 1963).
32. V. I. LENIN, *Imperialism, the Highest Stage of Capitalism* (first published, 1917) (International Publishers, 1939).
33. ———, *State and Revolution* (first published 1918) (International Publishers, 1939).
34. ———, "The State" (lecture delivered at Sverdlov University, July 11, 1919), *V. I. Lenin Collected Works*, Vol. 29 (Moscow: Progress Publishers, 1965).
35. HENRY M. LEVIN, et al., *Schools and Inequality* (Urban Coalition, 1969).
36. HARRY MAGDOFF, *The Age of Imperialism* (Monthly Review Press, 1969).
37. ERNEST MANDEL, *Marxist Economic Theory* (2 vols. Pearce, Brian, trans.) (Monthly Review Press, 1968).
38. KARL MARX, *Capital* (3 vols.; first English edition, 1887) (International Publishers, 1967).
39. ———, *Economic and Philosophic Manuscripts of 1844* (International Publishers, 1964).
40. KARL MARX AND FREDERICK ENGELS, *The Communist Manifesto* (first published 1848) (International Publishers, 1948).
41. J. E. MEADE, *Efficiency, Equality and the Ownership of Property* (Harvard Univ. Press, 1965).
42. ROBERT K. MERTON, ed., *Reader in Bureaucracy* (The Free Press, 1952).
43. STEPHAN MICHELSON, "School Resources and Equal Protection," *Inequality in Education*, Dec., 1969.
44. HERMAN P. MILLER, *Rich Man Poor Man* (New American Library, 1964).
45. C. WRIGHT MILLS, *The Power Elite* (Oxford Univ. Press, 1956).
46. JULIET MITCHELL, "Women: The Longest Revolution," *New Left Rev.*, Nov.-Dec., 1966. (Available as a pamphlet from New England Free Press, Boston, Mass.)
47. SCOTT NEARING AND JOSEPH FREEMAN, *Dollar Diplomacy* (Monthly Review Press, 1966).
48. TALCOTT PARSONS, "The School Class as a Social System: Some of Its Functions in American Society," *Harvard Educa. Rev.*, Fall, 1959.
49. KARL POLANYI, *The Great Transformation* (Beacon Press, 1957).
50. JOSEPH A. SCHUMPETER, *The Theory of Economic Development* (Oxford Univ. Press, 1951).
51. PATRICIA C. SEXTON, *Education and Income* (Viking Press, 1961).
52. ROBERT SOLOW, "The New Industrial State or Son of Affluence," *Public Interest*, Spring, 1968.
53. STERLING D. SPERO AND ABRAM L. HARRIS, *The Black Worker* (Atheneum, 1968).
54. PAUL M. SWEETZ, *The Present as History* (Monthly Review Press, 1953).
55. ———, *The Theory of Capitalist Development* (Monthly Review Press, 1942).
56. PAUL M. SWEETZ AND HARRY MAGDOFF, "The Multinational Corporation," *Monthly Rev.* (2 parts), Oct., Nov., 1969.
57. E. P. THOMPSON, *The Making of the English Working Class* (Random House, 1963).
58. U.S. CONGRESS, Joint Economic Committee, Subcommittee on Economy in Government, "The Economics of Military Procurement" (U.S. Government Printing Office, Washington, D.C., 1969).
59. MURRAY L. WEIDENBAUM, "Arms and the American Economy: A Domestic Convergence Hypothesis," *A.E.R.*, May, 1968.
60. RANDALL WEISS, "The Effects of Education on the Earnings of Blacks and Whites" (forthcoming in *Q.J.E.*, 1970).
61. WILLIAM A. WILLIAMS, *The Great Evasion* (Quadrangle Books, 1964).

ECONOMIC EFFICIENCY AND THE DISTRIBUTION OF BENEFITS FROM COLLEGE INSTRUCTION*

By W. LEE HANSEN, ALLEN C. KELLEY, and BURTON A. WEISBROD

University of Wisconsin

Economic efficiency implies an equating, at the margin, of benefits and costs. In this paper we explore a concept of "efficiency" which is broader than the usual framework and which applies to commodities and services produced and distributed largely outside the private, profit maximizing sector. An assessment of the economic efficiency of producing such a commodity requires the determination of its outputs and the valuation or weighting of these outputs.¹ Our principal point is that these weights, in turn, depend on who receives the outputs; thus, distributional issues are at the heart of economic efficiency studies involving a wide range of activities undertaken in the governmental and private, nonprofit sectors.² One of these activities—the production and distribution of college instruction in economics—illustrates well the significance of this particular approach to the analysis of economic efficiency.

We argue that in analyzing the economic efficiency of instruction, distributional issues—that is, who receives the benefits—should be considered explicitly; if not, they will necessarily be considered implicitly. The pervasive failure to include distributional issues in efficiency studies suggests an excessively narrow concept of efficiency. This is particularly inappropriate in evaluating instruction, since in education, as in most services, decisions regarding what to teach and how to teach have a strong influence on who receives the benefits.³

I

Total benefits from instruction are a function of the amounts gained by each student and of the values of each amount; these values vary among different students or types of students. Therefore,

* The authors are grateful to their colleagues, R. Andreano, F. Golladay, R. Lampman, and E. Smolensky, for comments on an earlier draft.

¹ We assume throughout that output is measured in incremental, units-added terms.

² This general point has been developed in detail by B. A. Weisbrod [17]. In his study the concept of efficiency which encompasses the traditional view of efficiency as well as the distribution of output is termed "grand efficiency."

³ This is true of most services because of the requirement that the consumer be present at the time and place of production (consumption) of the service.

aggregate benefits are a function of how the outputs are distributed among students. Symbolically,⁴ the marginal benefits, B_k , from resources employed in an instructional approach (technique and/or course content) k is:

$$(1) \quad B_k = \sum_{j=1}^n \frac{\partial q_j}{\partial k} \frac{\partial b_j}{\partial q_j},$$

where

q_j = the quantity of output produced by the input mix k and received by student j ; and
 b_j = the value of benefits (output) accruing to student j .

Of the two partial derivatives the first is the marginal physical product of input (or input mix) k for student j , and the second indicates the valuation of the marginal product.

As expression (1) reveals, the importance (both in quantity and in value) of any particular form of output may vary with the type of student recipient. For a student planning graduate work certain course outputs may have great value, whereas these same outputs may be of slight value to the student who plans to continue no further in the field. Similarly, instruction about behavior of the stock market may contribute greatly to the knowledge of persons from disadvantaged backgrounds while adding little to the knowledge of other students.

Most studies of instructional efficiency or studies which appraise the merits of particular teaching techniques have not estimated—nor have they even considered—the impact of alternative teaching approaches (input mixes) on the distribution of outputs among students. Neither have they examined the possibility that the value of outputs varies according to the distribution of the output. In terms of the model, the subscript j has been entirely disregarded. The assumption implicit in such a simplification is either that students are a homogeneous group—each student receives the same amounts of outputs from a given course and the outputs have the same value

⁴ In practice, the summation will be over groups of students who possess roughly similar and "relevant" attributes. Parenthetically, it might be noted that if $j=1$, expression (1) represents the value of an input used to produce a private good.

for each student—or that students should be treated as if they were a homogeneous group. Neither assumption seems reasonable, for rather obvious reasons.

This discussion has assumed implicitly that it is students' valuations that should count. There are many who feel, however, that valuations should be made by other groups, such as the faculty, the college administration, the taxpayers, or particular groups of students. Our general point—that the magnitude and distribution of benefits depends on the values assigned—is general and is invariant to whose values are applied.

In the following pages we discuss how instructional techniques and approaches influence the distribution of total output and the value of that output and how the distribution of output relates to a broad concept of economic efficiency.⁵

II

Consensus on course content or instructional technique has not been reached in economics or in general. The literature abounds with arguments for or against the "citizenship" focus [4] [5] [16] or the preparation of potential graduate students [13]. At another level we find some economists calling for a heavy mathematical orientation [2] (though whether as a "means" or as an "end" is not always clear); others advocate a decision-making framework [3]; and still others argue the merits of shifting emphasis from macro- to microeconomics [15]. And recently there has been much discussion of the impact of new instructional

⁵ Although much of college instruction is outside the public sector, our framework is relevant to the theory of the public sector which permits analytically separable allocation and distribution branches. If the allocation branch decisions affect the distribution of welfare, then it is not possible to determine what and how much of some commodity to produce unless costless lump-sum redistributions are possible, or unless the welfare function—involving the value-weights referred to in expression (1)—is known [10] [17]. The formulation in expression (1) permits the quantity of output to vary among consumers. This is by contrast with some of the literature on the pure theory of public goods, in which it is assumed that all consumers benefit equally in the sense of receiving equal quantities of output of the public good [14]. It is clear, of course, that even if all did receive equal quantities of output, the individual's valuations could differ greatly. Instruction does have a considerable public-good element; thus we might well have couched our argument simply in terms of differences in consumers' (students') valuations rather than in terms of differences both in their valuations and in the quantities of outputs. The difference in approach, however, is not substantive. Decisions will depend on the products of quantities and values—as expression (1) indicates—and our point is that these products are likely to vary significantly with the choice of course approach.

techniques, among them programmed instruction [1] [8], television [9] [12], and TIPS (Teaching Information Processing System) [6].

In all of these illustrations, either course content or instructional technique is being considered explicitly. But the distribution of benefits among types of student clientele is also very much at stake, and differences in judgments as to whom it is most "important" to teach may be at the root of much of the controversy. The issue of importance or values is reflected in the right-hand derivative ($\partial b_i / \partial q_j$) in expression (1), above.

Even if the importance of benefiting all types of students is equal or is assigned to be equal, the actual benefits (the left-hand derivative, $(\partial q_j / \partial k_i$ in expression (1)) are not likely to be equal. It seems intuitively clear that different types of students—as defined by previous academic performance, desire for theoretical rigor, degree of social concern, family background, and the like—will benefit differentially, depending upon course content and instructional technique. A highly theoretical and mathematical formulation in the basic economics course may provide the largest benefits for students already thinking seriously about graduate work in economics, whereas a course focusing on a less formal treatment of contemporary economic problems (and employing a similar instructional technique)⁶ may benefit most those students seeking a "general" education. Still another choice of course content is likely to be most beneficial to pre-law students. Given the heterogeneity of the enrolled students and the difficulty of offering simultaneously a variety of course contents, some students are certain to receive larger outputs⁷ than will others. Yet little or no evidence exists on the strength of the linkage between the distribution of outputs and the choice of course content.⁸

What has just been said about course content also applies to instructional technique. The traditional theory of production assumes that the

⁶ This qualification is intended to hold instructional technique constant in the present discussion so as to concentrate attention on course content.

⁷ Important problems exist with respect to the definition and measurement of "outputs"—they may well be multidimensional—but these are outside the scope of this paper.

⁸ In practice it may be difficult to separate the quantity of output from its value, and, indeed, the two may be interrelated. For example, if the consumer places a high value on increments of a particular type of course content, then his attitude toward learning may be "better," with the result that he will receive a larger quantity of output. This would make expression (1) more complex, for $q_j/k_i = f(b_j/q_j)$. An exploration of this complexity is beyond the scope of our present efforts.

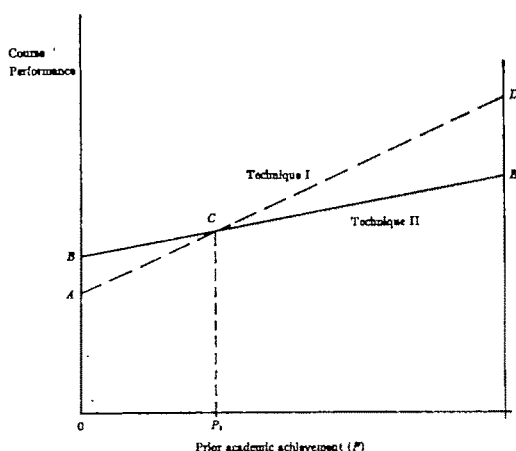


FIGURE 1

choice of production technique is a decision separable from the decision regarding the distribution of output. This assumption seems to dominate much of the evaluative research in economics education, for little effort has been devoted to finding out which kinds of students benefit by how much from the use of different instructional techniques; e.g., television and programmed instruction.⁹ The available evidence often shows whether the new technique is an improvement in the sense that students attain higher average scores than they do in courses using conventional methods. But the degree to which the new technique alters the distribution of performance is a subject scarcely ever raised.

The differential effectiveness of a given instructional technique for different groups of students and the differential value of a unit of "effectiveness" for different groups of students are illustrated in Figure 1. In this example, students are arrayed according to prior academic achievement, as measured on the X axis by a test given prior to beginning a particular course of study. The Y axis measures performance on an appropriate test of accomplishment following completion of the particular course. The curves labeled I and II show course performance—using alternative instructional techniques on two comparable groups

of students—as a function of the type of student within each group. "Course performance" reflects use of a particular test instrument for measuring "outputs" of the course.

The crossing of the two curves indicates that Technique I is more successful for the "stronger" students—those with prior academic achievement above P —whereas Technique II is more successful for the weaker students. Assuming an equal number of students at each level of prior achievement, P , it is clear that the average student gained more from the use of Technique I (that is, area CDE exceeds area ABC). Thus, if one and only one of the two techniques is to be used, the preferred approach is certainly number I. Or is it?

True, Technique I gives a larger mean and, hence, aggregate level of course performance. But what of the value of that performance? Despite the apparent superiority of Technique I, it would actually be inferior to II if the function for valuing benefits attached sufficiently greater weight to a unit benefit when realized by students with poorer prior achievement (e.g., the "disadvantaged"?). How to establish the values assigned is not an easy matter. But it is at this stage, involving valuation, that normative judgments are blended with positive findings as to the effectiveness of alternative instructional techniques or course contents.

The efficiency of Techniques I and II depends, ultimately, on costs as well as benefits. The technique producing the largest value of gross benefits (output) is not the most efficient choice if its costs are enough greater so that the value of benefits, net of costs, is smaller than for some other instructional approach. Moreover, it may be the case that none of the alternatives is efficient; that is, perhaps none produces net benefits that are positive.

Two important implications of the framework captured in Figure 1 might be noted at this point. First, the possibility of intersecting lines has profound implications regarding an analysis of the literature on the appraisal of teaching approaches. In general, although many studies in the education and economics literature show no significant impact of an experimental teaching approach, the studies are incomplete. They do not distinguish between the case of zero impact for all students and the case of positive impacts for some students and roughly equal negative impacts for others. Indeed, as McKeachie emphasizes, in his review of the voluminous literature on college teaching: "One reason for the host of experimental comparisons resulting in non-significant differences may be simply that methods optimal for some

⁹ This is true for all subject-matter areas, not only economics, and at all levels of education, not merely collegiate. This gap in educational research has been pointed out by others. For example, Robert Locke, an executive of the McGraw-Hill Book Company, has forcefully argued, "... the ideal program should not only use the most appropriate media for each task or objective, but it should also offer alternative media to accommodate differences among learners in ability, experience, motivation, style, and rate" [7].

students are detrimental to the achievement of others" [11, p. 1157]. However, neither we nor McKeachie knows which of the two cases above is closer to the truth. The implication is that much of the prior research needs to be reworked.

Second, if there is a distributional impact of the teaching approach (i.e., if the benefit curves of the experimental and control approaches are non-parallel, whether or not they intersect), then empirical tests which omit these distributional effects will produce statistically biased results.

III

One of us has experimented with a new instructional technique in the principles course in economics at the University of Wisconsin, and some of the preliminary findings are relevant here. The details of this experiment are not important to the present discussion. What is interesting is that this experimental approach, while having a beneficial impact on all students, provided larger quantities of output (in terms of the left-hand derivative in expression (1) above) to some groups of students than to others.

Figure 2 shows that the experimental technique, by comparison with a standard lecture technique (used with an essentially randomized control group of students), produced positive amounts of output for students at every level of prior achievement.¹⁰ But it also shows that the largest outputs went to students with the lowest ACT scores.

The fact that the experimental technique dominates the standard technique may appear to suggest that our findings are uninteresting in the context of concern about economic efficiency. Since the experimental technique involves added cost, however, this is not so. Even with the observed dominance, two issues remain unresolved: (1) Are the benefits disclosed in Figure 2—the area $ABCD$ —large enough to warrant the costs? (2) If total cost is a constraint, so that the experimental technique can be provided to some students but not to all of them, to whom should it be provided?

The answer to the first question depends critically on how the outputs (benefits) are valued; i.e., on the welfare function. There is, in general, no unique value for a unit of benefit (a point of added test score), for the value is a function of who the beneficiary is.¹¹

¹⁰ In this illustration students are classified by prior level of academic achievement; clearly, however, other classifications may be appropriate; e.g., in terms of family background.

¹¹ It may also be a function of the amount of benefit

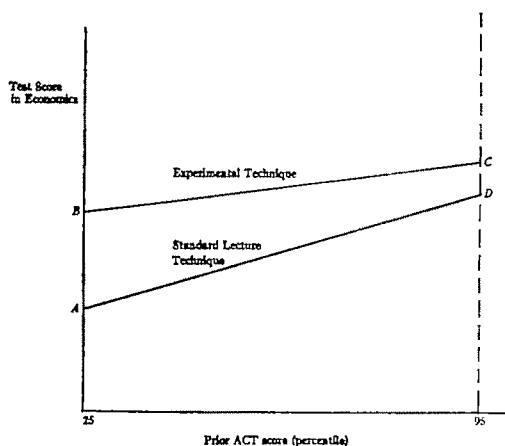


FIGURE 2

With regard to the second question, if the experimental technique is to be provided to, say, one-half of the students, which group should receive it? In other words, if total resources available are fixed, for which students will the value of output be a maximum?

The first thing to note is that Figure 2 cannot provide the answer. What it does show is that the program should be provided to the students with the lowest ACT scores if the objective is to maximize the measured quantity of output. But whether such an allocation would maximize the value of output is another issue. It will if the value of a unit increase in economics test score is a constant or decreasing function of prior ACT score, *ceteris paribus*.

If, however, there were a preference for helping "strong" students—a preference that may be reflected in the offering of "honors" courses—then it could be efficient to devote the added resources to students with high ACT scores.¹² All that is formally required to produce this result is that the value of output CD (Figure 2) exceed the value of AB , and that there be no sizable interaction effects among students.¹³ Since $AB = 5CD$

realized by any given beneficiary; that is, there may be decreasing, or perhaps increasing, marginal value with respect to added units of benefit to a particular student or group of students.

¹² It would be instructive to consider the forms of weighting functions that would justify providing large amounts of resources to particular subsets of students, as is done in honors courses and in special programs for the disadvantaged.

¹³ The latter assumption is necessary to insure that the separation of students into two or more homogeneous groups would not alter their respective class performance.

in the figure, an allocation of the experimental technique toward strong students is efficient if the value of additional economics output (as defined by the test used) be at least five times as great for a student at the 95th ACT percentile as it is for a student at the 25th percentile.¹⁴

IV

The arguments in this paper have a number of research implications. Granted that an efficient allocation of instructional resources requires comparisons of benefits with costs of alternative programs, more information is needed as to the magnitudes of the benefits. These, in turn, depend on the distribution of benefits, since the value of a given absolute increment in output of instruction (i.e., achievement) is specific to the recipient. Therefore, it is clear that we need to learn more about who benefits; that is, what types of students benefit and how much each benefits when various combinations of course content and instructional technique are used. With this information in hand, normative judgments can then be applied—as expression (1) above indicates—in order to estimate the value of benefits from any particular course content or instructional technique.

The following are some specific research suggestions:

1. Batteries of pre- and post-course evaluations should be developed which reflect well-articulated output goals. While this paper has not concentrated on a precise specification of outputs, the importance of such a specification is clear. The distribution of benefits (outputs) cannot be ascertained without a prior decision—either explicit or implicit—as to the definition of output. More tests measuring different types of outputs are needed.¹⁵

2. When regression techniques or other statistical procedures are used in evaluating teaching alternatives, interaction effects should be included so as to permit estimation of relationships between the instructional approach and a variety of student attributes (e.g., class, major, mathematics knowledge, family background). Parameters of the interaction effects—such terms as the application of television or programmed learning to students with and without calculus—will provide

information about the distributional impact of the teaching approach that is being studied.¹⁶ If students benefit differentially from alternative teaching approaches, then failure to account for interaction effects between the teaching approach and student attributes will result in a mis-specified statistical model, biased parameter estimates, and often uninterpretable results.

3. Finally, it is important to face up to the issue of delineating explicitly our normative criteria on what “should” be the distribution of outputs. We must place “values” on the outputs. Until these value weights are made explicit—and the task is not simple—they will continue to be implicit and, hence, not open to critical examination.

V. Summary

The central thesis of this paper is that production decisions in education on what and how to teach have distributional effects; as a result, distributional considerations should enter directly when making teaching decisions and evaluating these decisions.

The importance of this proposition and thus the validity of the inferences derived from it involve both factual and normative matters. If alternative teaching techniques do have differential impacts by type of student (the factual issue) and/or if social objectives dictate that it is more important to benefit some types of students than others (the normative issue), then the distributional consequences of selecting a teaching technique or course approach should receive explicit attention in benefit-cost analyses of production choices.

¹⁴ With respect to the pathbreaking research of Attiyeh, Bach, and Lumsden on the impact of programmed instruction in economics [1] [8] we are encouraged to learn that analysis is currently under way to investigate some of these types of interactions, and thus the distributional effects of programmed learning.

REFERENCES

1. R. E. ATTIEYEH, G. L. BACH, AND K. G. LUMSDEN, “The Efficiency of Programmed Learning in Teaching Economics: The Results of a Nationwide Experiment,” *A.E.R.*, May, 1969, pp. 217–23.
2. D. V. T. BEAR, “The Use of Mathematics in the Teaching of the Principles Course,” in Keith G. Lumsden, ed., *New Developments in the Teaching of Economics* (Englewood Cliffs, 1967).
3. J. G. C. BOOR, “Management Games,” in Keith G. Lumsden, ed., *New Developments in the Teaching of Economics*, *op. cit.*
4. H. S. ELLIS, “This is Economics,” *A.E.R.*, May, 1961, pp. 571–78.
5. R. FELS, “A New Test of Understanding in College Economics,” *A.E.R.*, May, 1967, pp. 660–66.

¹⁵ For this problem—with the cost level constant—only an index of relative values is needed, not absolute monetary values.

¹⁶ TUCE (Test of Understanding in College Economics) and TEU (Tests of Economic Understanding) represent an excellent beginning on this long-term project [5] [18].

6. A. C. KELLEY, "An Experiment with TIPS: A Computer-Aided Instructional System for Undergraduate Education," *A.E.R.*, May, 1968, pp. 446-57.
7. ROBERT W. LOCKE, "Educational Technology and the Educational Publisher," *Educational Technology*, Jan., 1968, pp. 14-16.
8. K. G. LUMSDEN, "The Effectiveness of Programmed Learning in Elementary Economics," *A.E.R.*, May, 1967, pp. 652-59.
9. CAMPBELL R. McCONNELL, "An Experiment with Television in the Elementary Course," *A.E.R.*, May, 1968, pp. 469-82.
10. MARTIN C. McGUIRE AND HENRY AARON, "Efficiency and Equity in the Optimal Supply of a Public Good," *Rev. of Econ. and Statis.*, Feb., 1969, pp. 31-39.
11. W. J. McKEACHIE, "Research on Teaching at the College and University Level," in N. L. Gage, ed., *Handbook of Research on Teaching* (Chicago, 1963).
12. DONALD W. PADEN, "Television as an Instrument of Instruction," in J. G. Lumsden, ed., *New Developments in the Teaching of Economics*, *op. cit.*
13. R. RUGGLES, "Relation of the Undergraduate Major to Graduate Economics," *A.E.R.*, May, 1962, pp. 483-89.
14. P. SAMUELSON, "The Pure Theory of Public Expenditure," *Rev. of Econ. and Statis.*, Nov., 1954, pp. 387-89.
15. G. J. STIGLER, "Elementary Economics Education," *A.E.R.*, May, 1963, pp. 653-59.
16. L. E. WAGNER, "Task Force to Classroom," *A.E.R.*, May, 1963, pp. 660-73.
17. B. A. WEISBROD, "Income Redistribution Effects and Benefit-Cost Analysis of Government Expenditure Programs," in Samuel B. Chase, Jr., ed., *Problems in Public Expenditure Analysis* (Brookings, 1968).
18. A. L. WELSH AND R. FELS, "Performance on the New Test of Understanding in College Economics," *A.E.R.*, May, 1969, pp. 224-38.

A RETROSPECTIVE LOOK AT UNDERGRADUATE ECONOMICS

By BEN W. LEWIS
The Ford Foundation

This occasion is something over which a man might well drool throughout a whole lifetime of anticipation. I am here by confirmed request to ramble through and reminisce about a lifetime; and, from the safe, secure vantage point of years of yesterdays, to recite to today's economists, for their benefit and redemption, their sins and shortcomings. "Tell us," you plead, "what it was like when the world and you were young. But more, tell us of our multiple re(trans)gressions and of what we ought to do that we are leaving undone. How have we strayed? How are we betraying our heritage?" This assignment, my friends, as you will someday appreciate if, like me, you are just luckier than hell, is the jackpot at the end of the rainbow.

True, there is a catch. I am invited to bask in my task—but only for twenty minutes. How much rich reminiscence and invaluable advice can be given wing in twenty minutes? This is not only to snatch back the Grail, it robs me of 60 percent of my fifty-minute birthright (quantitatively speaking). When the gavel strikes, my stage is to become a pumpkin, and I must be on my way wearing both slippers.

I am determined, however, to transmute this blight to bloom. If I am bereft of time, I am also relieved of obligation. No one can rightfully expect the accumulated insight of half a century to be both recounted and systematically supported in twenty minutes. I opt for recounting—stripped down. I shall recount and pronounce, and I renounce any obligation to defend my pronouncements. I shall offer no proof, no formulas, no hesitant, quavering hypotheses, and no ingratiating qualifications. Nothing but pronouncements in italics—sheer wisdom, topless truth! Let me add that nothing I shall say is confidential or restricted.

I

One moment while I lower the backdrop—the years I knew in my slender youth, the 1920's, the years of normalcy, the decade which ended in the Great Depression, the years 15–5 B.K.

Economics in the 1920's! The economy of which we taught was neat, clean-shaven, well-mannered, virile, self-disciplined, and rightfully confident. Its only markets were perfect markets (certainly perfect for us teachers), populated by single-

minded, uncomplicated economic men, compelled automatically and unerringly by irresistible urges to economic choices which were demonstrably correct for each and for all—in the long run. Undisturbed by aphorisms yet unuttered, we lived and reveled "in the long run." *Ceteris paribus*, in substantial doses, was our anodyne and constant comfort.

Competition was razor edged and all-pervasive—pure in mind and strong in body. We recognized only one degree of monopoly, and this we treated as an exception to be brushed off into a footnote located, in those golden days, at the bottom of the page, where footnotes belong.

Our forms of communication were, of course, primitive. The language of our discourse was written and spoken English. We used words—the kinds of sounds I am uttering now—extended into full sentences, in professional conversation and address; and words and sentences were to be found in abundance in our textbooks and journals. Their use was not surreptitious; it was open and unashamed, calling neither for explanation nor apology. The few graphs we employed were gentle and unobtrusive and had as their purpose to explain and illuminate the text—by today's practice, of course, an amusing transposition of horse and carriage. Economic models were unheard of; any models in which we may have had an interest were not permitted in the classroom.

Ours was an economy of principles, law and order. The "law of self-interest" and the "law of supply and demand" governed our micro concerns, and our macro concerns were few and slight. Say's law insured us against anything more troublesome than exceptional, transitional unemployment, and the Salvation Army was in the wings to alleviate any undeserved individual hardship. The role of taxes was to raise public revenues required to protect us against disorder and aggression; the role of money was to serve as a medium of exchange; the money was kept in order by Gresham's law. No one dreamed of questioning the constitutionality of our laws, and their enforcement was not a problem; they were self-enforcing.

The 1920's went out with a resounding thud. The Great Depression was a shaker. It shattered both the real world and our world. It had its redeeming features: it increased enrollments in

economics courses and it produced the first mass demand for college economists in business and government. Moonlighting was prestigious, not routine. Today's economists will never know how much their discipline owes to the New Deal for requiring dozens of their forebears, propelled into operational situations, to face up, "yes" or "no" to policy issues which they had been accustomed, at most, only to mull.

The depression opened the door to Keynes—another shaker, and a real mover. But, no one thought to close and bolt the door, and (I have steeled myself not to say "unfortunately") through the open portal came a couple of mathematicians. They stayed to dinner. They, too, were movers—they moved in. They multiplied (what else?); order appeared again—but this time it was numerical order; my old world was gone; economics began its frantic scrambling for recognition as an exact science. Economics without mathematics was no longer economics. The body of texts and tracts took on the appearance of appendices—ruptured.

I must make myself believe that my world was well lost. I cannot but admire an array of elegant skills, first joined to and then engulfing economics, that can measure the immeasurable, quantify that which cannot be quantified, replace dull words and vague discussions with fascinating symbols, strange devices, and instant answers to unanswerable problems—which knows no error, only deviations—and even these to be scooped up and run through again and again until, exhausted, they become absorbed.

I must make myself believe in mathematical economics, even though I will never understand it. It has made a real home for itself within the spacious reaches of our discipline—a towering, imposing, and ingeniously functional home with all of the charm of an oversize filing cabinet. But, brooding in my tiny, windowless room in the basement, I could wish that the structure didn't crowd the property line quite so closely on all four sides.

II

You will have gathered that I am less than completely charmed by today's economics, and that I indulge myself in an occasional flash of doubt about some of its manifestations. At my age, my attitude is to be expected; if it had not crept upon me along the way, I should have felt compelled to acquire it by forced feeding. And it is, of course, of no moment.

But, I do have a grievance which I believe is of moment. The real burden of the unhappiness

which I propose to unload here has to do with what we are doing and not doing with economics—and to economics—in the undergraduate classroom, particularly in the introductory course. I shall state my central proposition in terms it deserves: duty and sin. Economists engage in a variety of professional activities. Categorically, the most important of these, so overwhelmingly important that we cannot regard it as less than our greatest duty with no runner-up in sight is (a) to teach (b) the introductory course (c) as it should be taught to reflect fully our responsibility as a profession. Our greatest sin lies in our failure generally to accept and discharge this duty in the spirit and measure demanded by its urgency and consequence.

I shall not bother to argue here whether our posture towards the introductory course reflects cold calculation or aggravated indifference. Our apathy could not be deeper if it were, indeed, studied; it could scarcely be more widespread among us if we had taken a blood oath to effect and sustain it. The simple fact is that by and large, across and throughout our profession, the teaching of the introductory course, apart from the logistical problems it poses, is the very least of our concerns. We teach this course with our left hand, in our pocket. Exceptions are sporadic and short-lived. I am happy to except from this characterization the heroic efforts represented by this session's papers and other exercises promoted by the A.E.A. and the Joint Council. These reflect the birth pains of a systematic attack. In all honesty, however, any rift in the clouds that has yet appeared admits little light and no more than a teasing waft of warmth. It is far from clear that the pains augur more than a stillbirth.

Let me translate my evangelical message into native dialect. Economics is the study of the processes and criteria by which man selects the uses to which his limited resources are to be put. We accept as elementary that a use of higher importance or value is always to be preferred to uses of a lower order. Yet, in an area where we might be expected to be particularly sensitive—the allocation of our own valuable resources among competing professional activities—we do conspicuous violence to our own prime precept. It isn't that man-hours, apparatus and textbooks and bluebooks aren't poured into the introductory course. But what of their volume in relation to the job to be done, and what of their quality and the depth of the purpose which they represent? Thousands of us have taken our turn in teaching the course as part of our initiation into the mysteries of our profession while we were working off our

doctoral requirements, but at what cost in sacrifice of living students! Who among us has really had his heart in this enterprise, and for how long? How many choose it as a significant component of a career?

The case for giving top priority to the teaching of introductory economics is simple and forthright. You believe it or you don't. I do. It runs like this: Our discipline is important. An understanding of economics by all who participate is vital to the functioning and preservation of a self-governing society. Increasingly the character and quality of our society is being determined by collective economic decision making. The sheer fact is that the public understanding of economics demanded by this condition has never been realized and is today conspicuously in short supply. I cite the quality of the current national discourse on tax reform and inflation as a striking example. Certainly, not everyone in our society needs to be a computer-card-carrying economist; but I insist that it is of great importance to every adult member of our society that he and every other member shall have an elementary, systematic understanding of the economic underpinning of our lives together—the nature of an economy (all economies), and of the processes and institutions by which economizing takes place. My plea is not for fine-spun sophistication; it is for awareness and simple, basic economic understanding—for everyone. Provision of the means to insure such understanding fits—and bulges—all of the specifications of a public good.

Introductory economics in the colleges holds the key. It is only from this course that any hope is to be had for widespread, effective teaching of economics in the schools; and any significant extension of economic understanding in our society depends upon the further and deeper involvement of the schools. And it is only in the introductory course in the colleges that the great bulk of those who are to become college graduates have an opportunity for systematic exposure to economics, economies, and economizing.

Unfortunately, indeed tragically, far too few college students avail themselves of this alluring opportunity, and a major reason for this is that a large number (a "statistically significant" number) of those who have been conned into the course, or who have dropped in because they were cold or tired, have emerged disillusioned, still tired and noticeably cool. And bad news travels fast.

The news is bad for the introductory course, for our profession, and for our society. Because I have a stake in all three, I want my profession to

give more of its positive concern and its active, professional attention to the business of teaching introductory economics, and I mean business, not talk and not time serving; I mean coats-off, rolled-up-sleeves, dirty-hands, driving business. The introductory course is worth all the concerned attention we can give it, for what it can contribute to the rational, enlightened conduct of our society's economic affairs. It is worth everything we can give it to make it an attractive option for every student in the entire college community. I am quite shameless about this. We have something to sell that everyone ought to buy. We owe it to them and to ourselves to constitute and package, and everlastingly to reconstitute and repackage our product so that everyone, in fact, will buy.

I know the stock reaction; we all know it. What do you mean, we have too few students? We have, more now than we can properly admit, brand, bed down, grade, certify, and ship. We're overworked, we have no time for research and writing; we are mightily put upon. I admit all of this—and deny its alleged import. We may be pressing against the attention and the facilities which we are presently disposed to allocate to this task; undoubtedly we are annoyed by this yapping nuisance because the overwhelming bulk of our interest lies elsewhere, but we are not devoting anything approaching an optimum proportion of our energies and resources to the introductory course. I have spoken of this earlier, and I shall return to it; let me content myself here with the observation that it would be exhilarating to be confronted with logistical problems growing from an overwhelming surge of voluntary enrollments in introductory economics courses which were believed to be too good to be missed.

III

I have no chart or flight plan for the construction and conduct of *THE introductory-course-too-good-to-be-missed*. I have some ideas—all good, but none unchallengeable—about such courses. There is no single course suited to all times and all places; there is none suited to all times in any place, or to all times for any teacher. But, everywhere and eternally there is one irreducible minimum requirement for even a decent course: it must bear the bench mark of a concerned, deeply involved teacher, teaching at his level best.

Any course, superior, decent, or indifferent, is personal. It cannot be packaged and peddled as a proprietary product. Each of us must write his own prescription for his own course, and sign it. Your course is your course, just as my course is

mine. There is no escape; there should be none. This is far from saying that we cannot learn from each other, or that a man's course is a walled, bastioned castle, built for his protection. It is not his haven. None of us, if he is in the business, will find anything in his course but sweat, tears, and, hopefully, excitement—never comfort, and never more than scattered spots of satisfaction. If he finds refuge in his course, he may be on the staff and in the catalogue, but he is out of business. But, arena or refuge, it is his course; it can never be other than what he makes it and is willing to live with.

The essence of the experience is experimentation—constant experimentation, constant rethinking and constant trial and, of course, much error. This, I repeat, is the essence. The format of the course, its material and its emphases should be shaken up and repatterned year after year without cease—to shake up and rethread the mind of the shaker, and to demonstrate to students that learning is for everyone, not least a restless teacher. There is, of course, the off-chance that experimentation might, just possibly, lay a pearl. But if it does, the end of the road is still over the hill. Today's pearl is tomorrow's egg.

Parenthetically, I should like to record the judgment that the great promise of the present program of the A.E.A. Committee on Economic Education to improve teaching lies not in specification, but in what it may contribute to restlessness and experimentation.

In stating why the introductory course is our highest priority responsibility I have, necessarily, tipped my hand as to its orientation. I shall restate but not belabor my thesis. The introductory course should be policy-oriented, it should be called with deep conviction, "Introduction to Political Economy," and its conduct and thrust should redeem the promise of its title. The student—our student—should be lured into, instructed, informed, and involved intimately in, and made, as completely as may be, a living part of the total evolving political economy in which he, whoever he is, will live the rest of his life.

I am aware that the public policy problems which press upon us are not exclusively economic in content, and I trust I am quite clear that economics, either alone or in combination with other disciplines, holds out no promise of single right answers to our problems as distinct from all other answers which are wrong. But, there is certainly no major public issue, and no minor one comes to mind, which is not shot through with economic considerations; and economics does hold the promise of answers arrived at with a realiza-

tion of costs and consequences. I am firm in my conviction that if the economic policy choices we are driven to make are not supported broadly and firmly by the economic understanding of our people, the life we know will disintegrate. We will deserve to lose it.

Economies do not happen; they result from active and passive choice; workable economies are those that are made to work by combinations of individual and collective decisions. Running an economy—living and making a living together—in this crowded, complicated, confused century is proving to be a tough assignment. It has still to be demonstrated that we are up to the job. And time is wasting. Time is wasting, too, for the introductory course. It should be policy-oriented; it should be planned, packed, and pickled in economic public policy. Government, in its collective economizing role, should be woven into the presentation as an integral part of the economy. It is important for students to realize before they leave our loving care that collective economizing, exercised through the processes of their government, is no less "economic" than economizing by the market.

For the purposes which I believe should control, I would (to no one's surprise) build my course around the framework of "The Economic Problem." I would start with, return frequently to, and end on the strong notes of what, how, and for whom. The scheme and processes of economizing (both individual and collective, and the nature and essential tasks of an economy—all economies) would be laid out and nailed down. The great recurrent theme—the message—would be opportunity cost.

The scene would be "our economy," as it has been, as it is, and as it might be. The sets would be selected from the dozens of sectors and problems readily at hand, and I would mix them up and revise the selection frequently. I have seen too many popular areas and problems rise, flourish, gain ascendance, and then disappear without a trace to expect any of them to be here tomorrow. Fads are fun, but they are not for marrying. My problems would be drawn from markets of varying configurations and complexions and from the expanding field of government economic policy and action, and they would be recognizable by any practicing economist as both micro and macro in character.

Let me digress briefly here to give vent to what even I must accept as an idiosyncrasy. I would use the concept of opportunity cost quite consciously to bridge the chasm which we have insisted on excavating between micro- and macroeconomics.

I understand that micro is not macro, that individual choice is not collective choice, and that control of individuals by the market is not on all fours with control of the market (and its individuals) by government. So what? Are these distinctions imperatives? Precisely what is their economizing significance? I suggest that whether we are talking about individual investment in a canning factory or tax policy to control inflation or encourage growth, we are talking about choice in the use of scarce, valuable resources to accomplish what society finds acceptable, at the cost of other goods or ends foregone.

And while I am on this kick, let me air another, not unrelated, foible: a pox on flows! Money, capital, labor, trade do not flow. Individuals scent the air or feel its movement against an up-raised moistened finger; their nostrils quiver; they decide; they act. They do not flow: even as aggregates they do not flow. We economists may be architects and, sometimes, carpenters of a sort, but the mysteries of cosmic plumbing still elude us. The point is scarcely earth shaking, but neither is it inconsequential. It would be just as well if it were made quite clear to our students that even macroeconomizing involves considered actions by individuals—that economies are not to be turned on and off by fingering a faucet.

The introductory course—"Introduction to Political Economy"—as I conceive it, is the introductory course for all students, for all seasons—for the once-only or terminal student, for pre-majors (committed or still-to-be-seduced), for pre-engineer, pre-law, pre-business, pre-education students and for casual shoppers—for all college students except for those who can establish by examination that they have already been introduced. It should be taught as though it were the students' terminal course in economics, and (here I am probing, but on the basis of some revealing experience) as though all its students were prospective teachers. Try it sometime, and take your students with you.

The almost irresistible temptation to overload the table should be resisted. I suggest that the overall dimensions of the effort should be gauged by the needs of the terminal student, tempered by our best judgment of what he will take without gagging. This will provide quite adequately, as well, for those who are destined to take further work in economics—and it will leave enough undone to provide a rationalization for a light battery of advanced, undergraduate courses. And, might we dare to hope that somewhere in the intricate processes of the introductory course we might, just once chance upon the elusive secret of

imparting transferability, so that at least the most discerning of our students might be enabled to confront new problems with a working sense of "having been there before"?

Now, a comment about tools as a preoccupation of the introductory course. Regularly, almost dutifully, we recite the proposition that a major thrust of the introductory course should be to equip students with the tools of economic analysis. If this means leading students to approach economic problems as economists approach them and to think as economists think, the proposition is unexceptionable (at least if the prototype economist is, himself, unexceptionable!). But if, as I have reason to believe is too often the case, emphasis upon tools means how to construct and operate apparatus and appliances, the proposition is not persuasive. If we proceed on the assumption that students are crying, "Give us the tools and we will finish the job," we will find that the students and the course rather than the job are finished. Tools rarely invite or involve students not already preoccupied with mechanics, and there is little point in trying to equip students who aren't there. It is worth recalling that the preferred perch for Mark Hopkins was one end of a log, not a logarithm.

For most terminal students in the introductory course, elaborate tools will be costly excess baggage if they bother to pick them up, and a costly nuisance if they don't. Prospective majors can be toolled up as required in later courses. Since none of us wants to push students from the introductory course into the cold world of economic and political-economic affairs to make his way with nothing in his hands, I suggest that we might strike a sensible balance if we ask ourselves two questions: tools for what, and, is this tool necessary?

The usefulness of any tool for the basic purpose of the course should be clear and compelling. Before tools are exposed as tools they should be wrapped in economic problems that are intrinsically worth attending, and they should emerge unobtrusively from satisfying analysis of those problems. Students should come out of any exercise in mechanics with a sense of the meaning, not just the accuracy, of their calculations. Tools as such should never be given top billing, nor permitted to upstage the substance of the course.

I am informed that students now entering college are no longer frightened by tools, particularly by math, and that even before their beards are dry they are eager to employ their newly acquired manipulative skills. I have no reason to deny or deprecate this, just as I recognize and welcome

the better preparation many of them have received in, say, foreign languages and other disciplines—including English and economics. Some of these students, to be identified by examination, might be given a license and placed under observation in a special section; others might be kept in during recess or after class. In all cases without exception, however, I should still ask: tools for what, and, is this tool necessary? There is so very much, beyond tools, to be imparted and absorbed.

I am quite aware that the dichotomy I am posing is overstated, but my plea is for understanding as distinct from skills—understanding that finds its way through the irrelevant and the misleading to the heart of issues, and then captures the relevant and marshalls it economically and convincingly into the service of rational, purposive discourse and decision. I want students to feel economics, to like economics, and to live economics.

There is one demised tool, wielded by the teacher, which in its own unique, old-fashioned way, used to make a powerful contribution to economic understanding. I speak, in sorrow, of the passing of the essay examination. Its place can never be approached even remotely by true-false and multiple-choice examinations which, at best, are much more a test of the composer than of the student. We must not grieve, I know, over that which cannot be restored to us, but I must record what some of us feel deeply in our hearts. We and economic understanding have lost, not only a friend, but a staunch and resourceful ally. And I must shed one tear for the students of today and tomorrow, who, in their brave new world of slides, television, ear phones, transmitters, push

buttons, X's and checkmarks, will be forever denied the exquisite thrill of composing and committing to paper a complete, coherent sentence.

IV

I have been pretty prodigal here with my prejudices and precepts. I suppose that what I am really asking for is commitment—individual commitment. Every one of us is aware, perfectly aware, that there is an important job that must be done. With rare exception, every one of us avoids it. We all have our anesthetizing rationalizations, our convenient whipping boys; we are already overburdened with work of the only sort that carries professional recognition and advancement; the establishment, the board, the dean, the head are not interested; it's something no individual can do anything about; it's a job for the profession. Let me speak directly to these points.

Teaching economics is a great calling; it gets and will continue to get the recognition of a great calling in every instance where it is practiced as a great calling. No one else is responsible; we are not pawns or prisoners of any establishment, nor can we hide within the billowing folds of our profession. We—you and I—are our profession. And, for God's sake, who else can do the job? This is for us—us as individuals. Nothing happens in this world unless individuals make it happen. Remember? No flows! Any of us, alone, can affect the total outcome; and any of us, alone, can pick up company. But, of greater moment, individual outcomes in their own right can be of transcendent importance. Every one of us can look back upon one teacher who for him made all the difference. Well, all the difference is all I ask.

DISCUSSION

PAUL M. SWEETZ: The university is a part of the existing socioeconomic system—an increasingly important part—with specific functions, among which that mentioned by the authors (to socialize students as wage workers) is only one. Another is to inculcate into the student (a potential future leader, member of the intellectual elite, teacher of future students) precisely the point of view which underlies the economics curriculum; i.e., in the words of the paper, “acceptance of the *status quo* in social relations.” It follows—I’m afraid with iron logic—that the university is not going to accept, let alone promote, an alternative curriculum which begins by rejecting or calling into question all the basic institutions, relations, and results of the existing system.

It is true that under stress, such as that which developed on many campuses last year, the powers-that-be will make concessions—and often will promise more concessions. This is basically what Social Sciences 125 is: a concession or, more vulgarly, a sop. It is significant that the economics department refused to recognize it, and the academic authorities denied permission for it to abandon the grading system which the authors criticize so effectively in their paper. The chances that the experience with Social Sciences 125 will be allowed to form the basis of a new curriculum at Harvard, or anywhere else in the U.S. academic world, are pretty close to zero. And my own hunch is that it will be phased out as expeditiously as possible—maybe as and when the present staff, which developed the course, moves up or out.

None of this is said in criticism. Quite the contrary. It should be taken as praise. Any “radical” course or curriculum which would be acceptable on a permanent basis to Harvard—or any of the other major universities—would be, in my view, more than suspect. Its “radicalism” would be of the same kind as that of the British Labor Party, the European Social Democrats (and now we must add the official European Communist parties), African “socialism,” and similar enterprises in ideological confusionism. It follows that in saying that I do not believe this course will be long tolerated in the form presented in the paper under discussion, I am paying it the compliment of crediting its radicalism as genuine and solid.

Am I also implying that I think the whole thing is futile and a waste of time? By no means. Radical economics is, or at any rate can be, a very important part of a radical movement. Every adherent needs this kind of knowledge and understanding. All I am saying is that, unless I am very much mistaken, he is not going to get them in the formal offerings of America’s institutions of higher learning. As I said before, they have their responsibility which, despite all talk about the unfettered search for truth, etc., is in reality to support and service the capitalist system as it has taken shape in the United States.

If I am right, radical economics has to be taught by other than ordinary curricular means. Partly, of course, this can be done by radical economists in whatever officially recognized courses they teach. They can, and indeed if they are honest with themselves they must,

unsparingly criticize the institutions of capitalist society and the textbooks and other materials put out to rationalize and justify them. They do this in their own teaching and through supplementary readings, conferences, papers, etc. This, incidentally (or perhaps not so incidentally), is one reason why radical economists are not hired in large numbers by U.S. colleges and universities. A token representation is enough to prove devotion to liberal principles.

In addition to the efforts of individual radical teachers in their regular courses, it seems to me that more attention will have to be given to teaching radical economics in an extracurricular way. And it is here, I believe, that the paper by the staff of Social Sciences 125 can be most useful, and I certainly hope the entire text will be made fully available, not only through the *Papers and Proceedings* of this convention but also in a form or forms more readily available to the radical movement; e.g., as a pamphlet put out by the New England Free Press and/or the Radical Education Project.

On the whole I think the approach adopted by the authors is well suited to arouse the interest and enthusiasm of students disillusioned with their officially sanctioned economics (and other social science) courses. I think it should be possible to organize extracurricular programs with this as a beginning on many campuses. Variants will of course be tried and tested, and some may well be found to be improvements. But the basic content is not likely to be very different.

For my part I would like to see somewhat more emphasis placed on Marx and the Marxian tradition. Traditions can be valuable, educationally and otherwise, provided they are good traditions.

One final word: my experience with radical students in the last year or so engenders misgivings on one score. Many of them seem to think a radical education is not necessary or not important. They are largely activist-oriented, and some go so far as to assume implicitly, even if they do not openly state, that the revolution is coming in so near a future that they don’t have time to worry about getting an education under capitalism.

I wish they were right, but I think experience will show that we are in for a long, long period of crisis and struggle, in the course of which the radical movement will have to mature in strategic understanding and tactical wisdom. How that can come about without a good and effective radical education program, including economics as one of its central elements, I confess I do not pretend to understand.

PHILLIP SAUNDERS: I have been asked to comment primarily on the Hansen, Kelley, and Weisbrod paper. It is a pleasure to do so, for it is a very good paper. In emphasizing the determination of educational outputs and their valuation, they have put their fingers on two important and, as they emphasize, related issues.

For all of their attention to the problems of valuation and distribution, however, I would have welcomed more commentary on whose values are to be assigned to the right-hand derivative of their equation. They are correct in noting that their general point on the importance of valuation “is invariant to whose values are

applied," but it does seem to make an important difference if it is the valuations of the students or those of "the faculty, the college administration, the taxpayers, or particular groups of students" which are used. No one can quarrel with their final conclusion that "until these value weights are made explicit—and the task is not simple—they will continue to be implicit and hence, not open to critical examination." Indeed, it is precisely because "the task is not simple" that more attention to this point is needed. This is particularly true since the benefits of increasing most types of educational output very often accrue to one group while the cost of producing these benefits often falls on another group. To take only one obvious example, to which Hansen, Kelley, and Weisbrod referred in an earlier draft of their paper and which I wish they had left in, students are likely to prefer superoptimal rates of change in course content, instructional techniques, and related matters since they reap the bulk of the benefits, while the bulk of the costs of such change are borne by the faculty. On the other hand, and for the same reason, faculty and administration are likely to favor suboptimal rates of change—assuming, of course, that the valuations used in defining optimum are not exclusively those of one group or the other.

Turning to the left-hand derivative in the Hansen-Kelley-Weisbrod equation, the concept of value added is a tricky one in education—particularly if it is measured by before-and-after scores on objective tests with a finite number of questions. Before pursuing this point, however, let me add an aside that emphasis on value added is important. Most of the research in economic education to date (including, I might add, Figure 2 in the paper under discussion) has focused only on the level of final achievement which, of course, ignores the initial position of the student and, hence, value added. One reason for this is probably because of the lack of adequate test for before-and-after use—a situation which has, indeed, been improved with the publication of the Test of Economic Understanding for high school use and the Test of Understanding in College Economics for use in colleges and universities. But, one should take special note of the points that "more tests measuring different types of output are needed" and "the recently initiated project of the Committee on Economic Education of the American Economic Association and the Joint Council on Economic Education to develop a 'Question Bank' of validated multiple-choice items will further buttress research in economic education." (Those interested in more information on this latter project are urged to write Dr. Arthur L. Welsh, Director, College and University Program, Joint Council on Economic Education, 1212 Avenue of the Americas, New York, New York, 10036.)

Assuming adequate before-and-after tests, and given the importance of value added as the appropriate dependent variable, however, we are still left with the problem that simply counting the increase in test scores discriminates against those who begin with high scores. It is easier to improve your performance in absolute terms, for example, if you begin with a low score than if you begin with a high score on a test that has a ceiling imposed by a finite number of questions. Percentage changes don't help either, since a 5-point improvement

on a pre-test base of 30 would measure less than a 5-point improvement on a pre-test base of 40. The value-added measure that I think is the most appropriate one to use in these situations, therefore, is the one that was used by Simon Whitney; namely, the actual improvement expressed as a percentage of the difference between the before score and a perfect score on the test being used.¹ Alternatively, now that we have the Hansen-Kelley-Weisbrod model before us, one could value the points added over one range, say from 30–40, differently from the points added over another range, say 40–50. That is, use absolute increases in the left-hand derivative of their equation and then bring to bear the right-hand derivative. Given the range of weights that might be assigned to the right-hand derivative, however, the Whitney version of value added (recognized as an explicit combination of both derivatives in the H-K-W equation) may still have some merit.

Hansen, Kelley, and Weisbrod have performed a real service in focusing our attention on the distributional aspects of educational output. Some past research may, indeed, have to be reworked, and future research should now include interaction effects so as to permit estimation of relationships between the instruction involved and a variety of student attributes. In view of the very powerful influence of student ability found in most educational research to date, however, we may find that, after careful study, many of the interactions one can think of still turn out to be nonsignificant. I was fortunate enough, for example, to have received the Hansen-Kelley-Weisbrod paper in time to add an interaction variable similar to the one they used in Figure 2 to two studies I was in the process of completing. In both cases I got a result similar to the one they have plotted in Figure 2, but neither result was significant statistically. Until much more of this type of testing is done, however, one can only agree that "it is clear that we need to learn more about what types of students benefit, and how much each benefits when various combinations of course content and instructional technique are used."

Before closing, I cannot resist the opportunity to say a few words about the other two papers also presented here. In the case of Ben Lewis' paper it will be a very few words—three to be exact. "Amen, Brother, Amen."

I did not receive the Edwards and MacEwan paper in time to pursue thoroughly all of the references and arguments cited in their many footnotes, but in attempting to show how social problems such as alienation, imperialism, inequality, racism, etc., are "directly attributable" to the operation of the institutions of capitalism and the class divisions which those institutions spawn, it seems to me that they are hard put to account for these same problems (often in much more acute form) in other societies which cannot by any stretch of the imagination be called "capitalist."

One doesn't have to read Milovan Djilas to know that social classes are not peculiar to capitalist societies, and the tendency of those who run the state to use it to preserve their own interests doesn't strike me as unique. Nor does it seem that the "ideology of capitalism" is

¹ See Simon N. Whitney, "Measuring the Success of the Elementary Course," *A.E.R.*, Mar., 1960, pp. 159–69.

alone in placing "a high value on the rise of material output." Eisenhower did not lose his job over his country's sagging economic growth in the 1950's; Khrushchev did when the Soviet economy sagged in the early 1960's.¹ Turning to particular problems such as alienation, for example, it was not a capitalist society that erected the Berlin wall and posted sentries to shoot and kill the hapless individuals bold enough to risk their lives in fleeing the barren emptiness of their existence in a "workers' paradise." With regard to imperialism, it was not capitalist imperatives that sent tanks into Budapest in the 1950's and into Prague in the 1960's. Empirical evidence indicates that wage differentials and inequalities are greater in the Soviet Union than they are in the United States. Is racism and the subjugation of women worse under capitalism than it was under feudalism? Etc., etc.

Many of the problems that Edwards, MacEwan, and the staff of Social Sciences 125 are concerned with seem to me to be problems of industrialization—regardless of the form industrialization takes, capitalist or otherwise. And, in talking of "traditional, nonalienating work processes," it also seems to me that, like many others who are worried about the problems of industrialization, they tend to look at preindustrial or nonindustrial societies with rose-colored glasses. I wonder how many Athenian slaves and medieval serfs would rise from their graves and testify to the lack of alienation in their humane and self-fulfilling occupations.

Do not misunderstand me, the problems with which Edwards and MacEwan are concerned are real, and they should trouble all thoughtful men. But, in prescribing solutions for these problems, it is important to diagnose them correctly. Herbert Simon has noted: "You can have a crisis for two reasons: you can have a crisis because the world has changed and it threatens goals that you previously had; but you can also have a crisis because the world is unchanged, or it has changed too slowly and your aspirations have risen. If we have a crisis today, I think it is the latter kind, and it was Barbara Ward, I think, who christened it 'a revolution of rising aspirations.'"²

Thus I feel that Edwards and MacEwan have missed the point when they state: "the rise in concern regarding the problem of pollution is in no way surprising, nor should it be surprising that antipollution groups make headway only when the problems become severe." Objectively, there is less air and water pollution in Pittsburgh today than there was twenty years ago; yet active concern has never been greater. Indeed, it has led to the formation of a group whose aims I like as much

as its name: GASP—Group Against Smog and Pollution.

The same problem of expectations outrunning our ability to perform rather than the system deteriorating because of certain class imperatives also cuts across other points raised in the paper. While thus weak in diagnosis, the Edwards and MacEwan paper is also almost completely silent on explicit prescriptions for dealing with the problems they have raised. They note that they end their course by examining the answer to the question: "How can radical social change take place?" Yet they don't tell us what the answer is.

In conclusion, Harvard's Social Sciences 125 appears to be an exciting and an interesting course. If taught on the nonauthoritarian and (I assume) nondogmatic grounds the authors expound, it might be a lively supplement to a student's education after he or she has completed an introductory economics course like the one Ben Lewis has outlined. I do not, however, see it as a replacement for an introduction to economics, let alone as the basis for an entire economics curriculum. The analytical concepts outlined are more sociological than economic in nature, and as the analytical core for an entire curriculum they seem far too limited. Much of the present one-year course is already made up of case studies. Beyond simply adding more and more cases, the authors do not indicate what additional concepts in radical economics would be added to the analytical core already covered.

JOHN R. COLEMAN: My agreement with what Ben Lewis says is so strong that the temptation is to rush quickly into singing his paper's praises. But the paper by Richard Edwards, Arthur MacEwan, and the Harvard staff members in Social Sciences 125 raises so many questions in my mind that I resist the urge to speak of Lewis and open instead with some comments on the Harvard paper.

Radical economics may well be here to stay. It may even have an important contribution to make to the processes and uses of our discipline. But the Edwards-MacEwan paper doesn't prove it. It leaves us with the nagging question posed by so many of today's radical movements: how long are we to suspend belief?

I am ready to join in sharp criticism of many of today's economic institutions. I am ready to say that wise and humane men can surely do much better than we have done with all our resources. But how long must I take it on faith that the new left has better answers than the rest of us? This paper is fifty pages long in typescript. In those fifty pages, there is one paragraph only on the economic structure which is to replace the current and criticized one. The only hint of that world is too vague to help us: "In particular we examine how incentives, the control of the work process, and the process of distribution are organized in other societies (e.g., Cuba, China, Yugoslavia) and how they might be organized in a socialist society in the United States." That may be the stuff of which dreams are made, but I wonder if it won't wear thin a few months into the course.

Elsewhere we are told, for example, that income determination in our capitalist economy is a struggle. True. But is that not a characteristic of every economic man has yet devised? Is capitalism the problem, as

¹ Incidentally, on an analytical fine point, the equation of maximum output with maximum profit in the phrase "... process of yielding a maximum output (maximum profit) situation ..." is inconsistent with the conventional theory of the firm which shows that it is by restricting output (compared to the output of a perfectly competitive firm with identical cost curves) that the monopolist maximizes profit. This inconsistency is underlined since the authors previously emphasized "... the monopolistic conditions of business in the United States."

² Herbert A. Simon, "The Problems of Cities," *Carnegie Rev.*, Oct., 1969, p. 4.

these men suggest, or is it the more universal problem that there is only so much to go around at any time? This paper sees economic conflicts at the core of the most pressing issues ahead of us. And that is presented as something of a new finding, whereas I had thought many of us have been teaching that as a central organizing theme in our economics courses for generations. That is what we thought we meant when we talked about the allocation of scarce resources among competing ends. So what else is new in the radical economics?

I am troubled, too, by the description given by Edwards and MacEwan of elementary courses other than their own. Who among us, for example, teaches that collective bargaining is "predicated on the assumption that the struggle is one over the distribution of 'excess profits' "? Not a labor economist, I trust. And who teaches that the only issue for economists so far as Vietnam is concerned is "how it can be financed more efficiently"? No economist with an awareness of what that war is doing to so much we have believed in about ourselves and our nation, I trust.

The new course, insofar as we are told what it is here, appears then to be an updated indictment of industrialization more than of capitalism and a plea for something better but undefined. What does the fact that able young men wish to teach such a course and others to enroll in it say about what the rest of us have been doing?

I fear that it simply confirms much of what Ben Lewis is saying. It also confirms the fear I have expressed elsewhere that we have lost sight of values in our discipline or have treated them so lightly that it would have been better to ignore them completely. We are not asking, "Economics for what?" in a way that we must do if we are to have a fighting chance to hold the best young minds around us.

At their best, the radical economists are bringing legitimate, vital, and too long ignored questions into the classroom. Are we?

At our best, we are bringing disciplined thought into the classroom. Are the radical economists?

The challenge to us as teachers of economics has never been clearer. Yet there is little evidence that we are ready to respond. If we merely sneer at those who are loose with facts or who have as many inconsistencies in their work as we do, we will do so at our peril, not at theirs. We have a bigger task on our hands: to prove once again, to a more skeptical generation, that our discipline can illuminate aspects of the awesome choices ahead of this society and this world.

Taking Ben Lewis seriously would carry us a long way towards meeting that challenge. Here is a teacher whose bite is still worse than his bark and whose delightful style scarcely conceals some profound and not so delightful conclusions. The number of leaders in our profession who have kept a lively and fresh concern for the first undergraduate course is small. Add up all the Ben Lewises, the Leland Bachs, the Henry Villards, the Clair Wilcoxes, the Rendig Felses, the Paul Samuelsons and the Phillip Saunderses, and you still have only a handful of men. Yet the elementary course is where most of us learned our economics—though probably not as students, but as teachers of that course.

The key to Ben Lewis' influence is in his recognition that the elementary course must ultimately be the "bench mark of a concerned, deeply involved teacher, teaching at his level best." (That concern may also be the strength of the new radical economists.) Our contributions ultimately come down to concern plus discipline. Our professional training tells us that facts are important, and that they sometimes change over time. Our interest in the allocation of scarce resources must also tell us that it does make a difference what man chooses to do with his know-how.

Unfortunately our classroom teaching is all too often carried on in a sealed area from which any peer who might be helpful to us in striving for that level best for which Lewis pleads is excluded. It is well to experiment, as Lewis suggests. But experimentation's handmaiden must be evaluation, and we don't do that part well as long as we perpetuate the classroom-as-castle syndrome. Scientific and aspiring in so much of our work, we act otherwise when it comes to learning how to be better teachers.

The case for political economy as the central thrust of our course is well made by Lewis. Perhaps the truly radical way to begin to implement what he calls for is to start talking to our colleagues in the introductory government courses. That would have the side advantage of making some of our breast-beating about the state of teaching the elementary courses contagious.

If I miss any one note in the Lewis paper, it is an explicit plea for not overloading the course. A few basic ideas taught so well that clearer thinking among our students is the product should be enough for any of us. To do that would require the courage to face one's colleagues and to admit openly that such and such an "indispensable" topic has been dispensed with. Incidentally, it is just possible that these meetings of our Association do more harm than good if they give a further forum, through corridor conversations, to those who have a long list of things that must be "covered" in the basic course.

We can sacrifice coverage with ease if we have the sense to agree on a few central aims for the courses we teach and then have the courage to stick with these aims. The final exam can then be the testing ground for the fulfillment of our objectives. In twenty years of teaching the elementary course, I can only recall once—last year—when I gave a final examination that really worked well in giving perspective on the central ways of thinking for the course, in discriminating among students in their ability to use those ways of thought, and in proving enjoyable too.

The examination simply asked students to design a five-question macroeconomic literary test for United States congressmen and to show why each question might usefully separate the economically sophisticated from the naïve in Congress. The results were some fine papers. I felt the class saw what was important in the course. And I sensed their interest was alive enough that they would have enjoyed seeing their tests administered in Washington.

Ben Lewis has talked straight and hard in this paper. The entire profession is in his debt once more. When do we start paying off that debt? The place to begin is in the next elementary class we face.

INVITED DOCTORAL DISSERTATIONS

STOCHASTIC STABILITY OF A GENERAL EQUILIBRIUM MODEL*

By E. ROY WEINTRAUB
Rutgers University

I. Introduction

Ever since Walras and Pareto, the problem of general equilibrium has received attention from economic theorists. The ultimate concern of the investigations has been rather more concerned with the methodological questions than the policy implications; no one pretends that the set of equations one deals with describes reality. Instead, the viewpoint is taken that through study of the several assumptions and derived equations, the interactions of the various economic agents can be made logically clear.¹

Yet the formal models proposed to lay bare the complex network of the price system must, to meet our approval, pass several tests. First, such a model which seeks to elucidate the workings of a price system ought to possess an equilibrium; that is, there ought to exist some set of prices which equilibrate the myriad conflicting interests of the transactors. From studies associated with such names as Wald, Debreu, Arrow, McKenzie, and Gale, it has been learned that assumptions about the maximizing behavior of the agents, perfectly competitive organization of the markets, and various regularity conditions suffice to insure the existence of an equilibrium set of prices.

The second major test usually put to these general equilibrium models is stability analysis; that is, if one postulates a dynamic adjustment mechanism, will an initial nonequilibrium set of prices approach the equilibrium set? Work by Arrow, Hurwicz, Negishi, Hahn, McKenzie, and others has suggested (roughly) that if the competitive models are such that gross substitutability prevails throughout the system and prices move in the direction of excess demand, then the model is stable: over time the equilibrium prices will be approached.

* This paper is based on the author's Ph.D. dissertation submitted to the Faculty of the Univ. of Pennsylvania in May, 1969, under the supervision of Professor Lawrence R. Klein. The author wishes to acknowledge the helpful comments of Jerry Kazdan, Steven Turnovsky, Edwin Burmeister, Myron Myer, and, particularly, Lawrence Klein.

¹ R. Dorfman, P. A. Samuelson, and R. M. Solow, *Linear Programming and Economic Analysis* (McGraw-Hill, 1958), p. 351.

For one who believes in the fruitfulness of a general equilibrium theory, however, optimism is short lived. Both Scarf and Gale have shown clearly that there are quite reasonable competitive equilibrium models which are totally unstable. Thus the theorist can take one of two positions. He can argue that the assumptions of the stable models are sufficiently satisfying since they yield strong results, or he can argue that the assumptions of Scarf and Gale are not intrinsically unreasonable, so that if one wishes stability one might have to go outside the perfectly competitive framework.

One tentative approach directed to adjudicating the two positions was suggested in my dissertation. Heuristically, consider the following argument. Even in a competitive framework, the decisions that the various economic agents make are made under conditions of uncertainty; in all markets there are random forces at work. These random shocks will thus appear in each of the interrelated equations modeling the price system. If the stability or instability of this now stochastic competitive system can be determined, one may draw some inferences. In particular, if the random shocks are stabilizing, then by incorporating more realism in the model one has shown that the system is more stable. Those who feel that the perfectly competitive model is instructive are thus strengthened in this belief. Conversely, if the random shocks destabilize deterministically stable competitive models, then more credence must be lent to those who suggest that one ought to turn to imperfectly competitive models to seek stability.

Put briefly, the results support the latter view; that is, reasonable random shocks can destabilize the usual (stable) competitive equilibrium models. To demonstrate this, some formalization will be necessary, although the mathematical details are reserved for the appendix.

II. Assumptions

Consider a pure exchange economy such that:

1. The economy consists of n individuals with holdings of m commodities.
2. The total commodity stock is constant.

3. Each individual seeks to maximize his total utility through exchange.

4. The relative prices of the m goods are initially given in the market.

5. Individual demand for each good is a function of the prices of all goods.²

Let

p_j = price of j^{th} good
 \bar{x}_j = total stock of j^{th} good
 x_j = total demand for j^{th} good
 p_j^* = equilibrium price of j^{th} good,

and assume that

h) The demand functions are continuous and homogeneous of degree zero in prices and

w) Walras law holds: i.e. $\sum_j p_j(x_j - \bar{x}_j) = 0$.

The above, with nonnegativity assumptions on initial holdings and convexity assumptions on preferences, suffices to insure the existence of an equilibrium price vector p^* so that $x_j(p^*) = \bar{x}_j$ for all j . To further simplify matters, we also assume that equilibrium is unique (up to a scalar multiple).

To study the dynamics of the system, we shall assume a tatonnement mechanism:

$$t') \quad \frac{dp_j}{dt} = x_j(p) - \bar{x}_j \quad \text{for } j = 1, 2, \dots, m.$$

Setting $f_j(p) = x_j(p) - \bar{x}_j$ as the excess demand function we have, alternatively,

$$t) \quad \frac{dp_j}{dt} = f_j(p) \quad \text{for } j = 1, 2, \dots, m.$$

That is, the rate of price change is proportional to excess demand.

III. The Deterministic Case

The usual results can be presented as follows:

Definition: If $p(t) = (p_1(t), \dots, p_m(t))$ is any solution of t), then p^* is asymptotically stable if $\lim_{t \rightarrow \infty} p(t) = p^*$ for $p(0) > 0$; that is, for any given initial (positive) price configuration, through time those prices converge to equilibrium.

Theorem 1: If $\sum_j p_j^* f_j > 0$ and h) and w) and t) hold while f_j is continuously differentiable, then p^* is asymptotically stable.

In other words, the system is stable if both the usual assumptions hold and the equilibrium price weighted excess demands are strictly positive.

² These are the assumptions of Negishi's article, "The Stability of a Competitive Economy: A Survey Article," *Econometrica*, Oct., 1962.

The economic assumption that can guarantee this is gross substitutability.

Theorem 2: If $(\partial f_i / \partial p_j) \geq 0$ for $i \neq j$ then $\sum_j p_j^* f_j > 0$.³

Thus homogeneity, continuous differentiability, Walras law, and gross substitutability entail asymptotic stability of the exchange equilibrium with tatonnement dynamics.

Now if it can be shown that in the presence of random shocks these assumptions do not suffice to insure stability, we shall have demonstrated the destabilizing influence of random shocks.

IV. The Stochastic System

In its most general form, the equation which expresses the influence of random behavior on the dynamics of the system is

$$(1) \quad \frac{dp_j}{dt} = f_j(p, y(t)); \quad j = 1, 2, \dots, m.$$

Here $y(t)$ is a random function in the sense that, subject to certain regularity conditions, for every t , $y(t)$ is a random variable.⁴ That (1) is a legitimate generalization of (t) may be seen as follows. Suppose that an initial set of prices is called out in the market, and excess demand for good j is positive. A higher price for good j is thus announced. A consumer, however, may be uncertain about what to expect in the future. For if his expectations are formed statically, so that, to him, the expected price usually equals the current price, with uncertainty it is possible that expected price equals current price plus a random variable; that is, if the expected price is denoted by \hat{p}_j , the tatonnement could be expressed as

$$t'') \quad \frac{dp_j}{dt} = f_j(p, \hat{p}) \quad \text{and} \quad \hat{p} = p + y.$$

With uncertainty we might have

$$(2) \quad \frac{dp_j}{dt} = f_j(p, \hat{p}) \quad \text{and} \quad \hat{p} = p + y$$

for y a random variable. (2) is really in the form of (1).

Several questions arise quite naturally regarding (1). Since the excess demand function is now stochastic, the rate of change of prices, and hence the time path of prices, is a random function. What does an equilibrium mean in this case? Does sta-

³ Theorems 1 and 2 are the crux of the original proof in K. Arrow, H. D. Block, and L. Hurwicz, "On the Stability of a Competitive Equilibrium II," *Econometrica*, Jan., 1959.

⁴ See, for instance, A. M. Yaglom, *An Introduction to the Theory of Stationary Random Functions* (Prentice-Hall, 1962).

bility have any meaning? And if stability does have meaning, do the usual assumptions about the excess demand function insure that the stable equilibrium will be attained? While leaving details for the appendix, the analysis can be sketched intuitively.

V. Stochastic Stability

Before answering the stability questions, two problems about the interpretation of (1) arise immediately. First, what is the form of $y(t)$, the random disturbance or shock? Second, how is the shock incorporated into the excess demand function? To answer the first question, we can use here in continuous time the continuous analogue of the econometricians' discrete time errors: standard, normal completely uncorrelated random variables with constant spectral density. The continuous analogue is Gaussian white noise, a mean zero process with infinite variance;⁵ that is,

$$(3) \quad E[y(t)y(s)] = \delta(t - s)$$

where δ is the Dirac function.

Regarding the second question, one can remain neutral; that is, by writing (1) as

$$(4) \quad \frac{dp_i}{dt} = f_i(p) + \sigma_i(p)y(t); \quad i = 1, 2, \dots, m.$$

with σ_i continuously differentiable one has simply separated out the purely deterministic portion of the excess demand function. Obviously $f_i(p)$ will retain its usual properties. $\sigma_i(p) = 0$ represents the usual deterministic case; if $\sigma_i(p) = a$ constant, the shocks are additive, and if $\sigma_i(p) = cp_i$ for c a constant, the shocks could be termed multiplicative.⁶

The meaning of equilibrium for the stochastic system (4) can be dealt with quickly. The equilibrium price vector should be the same whether or not shocks in any form are present. The shocks affect the time path of prices, the dynamic adjustment, and not a price set determined from static considerations. The constant vector p^* retains all the properties of an equilibrium.

There are several problems, however, with the meaning of stability. In the usual case, asymptotic stability means that from any initial set of prices, prices eventually get close to the equilibrium set. Stability is thus akin to convergence.⁷ It is this

viewpoint which is useful for the present stochastic considerations. An arbitrary time path of prices, a solution to the system (4), is now a random function and as such is represented by a probability distribution at each point in time. The equilibrium motion is constant. One thus asks whether the stochastic motion converges to the equilibrium. But there are many notions of stochastic convergence: convergence in probability, convergence with probability one, convergence in the r th moment (where for $r=1$ we would have the expected value of the prices approach equilibrium), etc. Each variant of stochastic convergence yields a corresponding definition of stochastic stability; e.g., stability with probability one.

The final question posed at the end of Section IV, of course, is the crucial one: is the system (4) stochastically stable in any sense under the usual assumptions on the excess demand function which served to insure (deterministic) asymptotic stability? The following result has been established: *Theorem 3*: Consider the stochastic tatonnement given by equation (4). Assume f_i and σ_i are continuously differentiable, and that f_i satisfies h and w . Define $M = \min\{p_1^*, p_2^*, \dots, p_m^*\}$ and

$$v(p) = \frac{1}{2} \sum_{i=1}^m (p_i - p_i^*)^2.$$

It

$$\sum_{i=1}^m p_i^* f_i > \frac{1}{2} \sum_{i,j=1}^m \sigma_i(p) \sigma_j(p),$$

then $p(t) \rightarrow p^*$ with probability $1 - v(p)/M$.⁸

A proof of this as well as a rigorous set of stability definitions will be found in the appendix. Let us here interpret the theorem. The first main point is that in the deterministic case, stability was guaranteed if $\sum_i p_i^* f_i > 0$. In this stochastic case, not only does $\sum_i p_i^* f_i$ have to be positive, but it must be larger than a positive number which depends on the strength of the shocks, $\sigma_i(p)$. This is clearly seen if $\sigma_i(p) = a$ constant. In the usual case, gross substitutability guaranteed the positivity of $\sum_i p_i^* f_i$. Thus the theorem suggests that gross substitutability is not enough to insure $\sum_i p_i^* f_i > \frac{1}{2} \sum_{i,j} \sigma_i(p) \sigma_j(p)$ since the shocks are arbitrary, and hence gross substitutability does not suffice for the weakest forms of stochastic stability, stability in probability.

Second, if one is dealing with local stability, so that one starts close to equilibrium, $v(p)$ is quite

⁵ *Ibid.*, p. 64.

⁶ For a discussion of the difference between additive and multiplicative shocks in discrete time dynamic systems see S. J. Turnovsky, "Stochastic Stability of Short-run Market Equilibrium Under Variations in Supply," *Q.J.E.*, Nov., 1968.

⁷ See, e.g., J. E. Bertram and P. E. Sarachik, "Stability of Circuits with Randomly Time-Varying Parameters," *IRE Transactions, Supplement to Volume CT-6*, May, 1959.

⁸ This theorem is obtained as a corollary to the general theorem proved in Chap. 2 of H. J. Kushner, *Stochastic Stability and Control* (Academic Press, 1967). Kushner's general work is the basis for the analysis in this paper.

small in relation to M , and so the conclusion of the theorem would be that locally $p(t) \rightarrow p^*$ with probability one. (This is equivalent to proving a global probability one stochastic stability result for the linear approximation to the system.)

Third, if $\sigma_j(p) = 0$, which is the deterministic case, the theorem reduces to the usual result.

VI. Conclusion

In short, the mathematical formalization of the *gedankenexperiment* whereby the dynamic adjustment mechanism is perturbed by random shocks yields intuitively plausible results under certain conditions. The disturbances do in fact destabilize the usual competitive model and thus suggest that stabilizers must be sought elsewhere.

Lest this result be thought too nihilistic, it can be shown that the existence of a stock as well as a flow adjustment mechanism can in certain circumstances "damp" the random shocks, smoothing them sufficiently for stability to remain unaffected by disturbances. And it is a reasonable conjecture that if the random shocks were correlated from equation to equation in the dynamic system, this might even be a stabilizer to price movement.⁹ On the other hand, it can be shown that methods of forming expectations different from the static case considered here have no appreciable stabilizing influence on the stochastic system.¹⁰

MATHEMATICAL APPENDIX

Consider equation

$$(A1) \quad \frac{dp_j}{dt} = f_j(p, y(t))$$

where p_j, f_j , and $y(t)$ are as in Section V.

Definition 1: $p^*(t)$ is an equilibrium motion of (A1) if $f_j(p^*, y(t)) = 0$ for all j and $t > 0$ and for all random functions $y(t)$.

The various stability definitions are as follows: **Definition 2:** $p^*(t)$ is asymptotically stable in the mean (relative to H_0) if $\lim_{t \rightarrow \infty} E|p(t) - p^*(t)| = 0$ for $|p(0) - p^*(0)| < H_0$.

Definition 3: $p^*(t)$ is almost surely asymptotically stable (alternatively, asymptotically stable with

probability one) (relative to H_0) if $\lim_{t \rightarrow \infty} |p(t) - p^*(t)| = 0$ with probability one for $|p(0) - p^*(0)| < H_0$.

Definition 4: $p^*(t)$ is stable in probability if for all $\epsilon > 0$ and $q \in (0, 1)$ there exists a $\delta(\epsilon) > 0$ such that $Pr\{|p(t) - p^*(t)| < \epsilon\} > 1 - q$ for $|p(0) - p^*(0)| < \delta$.

It is easily proved that either asymptotic stability in the mean or almost sure asymptotic stability imply stability in probability.¹¹

As in Section V, the equation to be studied can be written as

$$(A2) \quad \frac{dp_j}{dt} = f_j(p_1, \dots, p_m) + \sigma_j(p_1, \dots, p_m)y(t) \quad j = 1, 2, \dots, m$$

where f_j and σ_j are continuously differentiable, $y(t)$ is white noise, and f_j satisfies h and w . As Kushner has shown, any solution of (A2) is an Ito process which is stationary Markovian, and continuous w.p.l.¹²

We now wish to prove

Theorem 3:¹³ Assume (A2) with f_j, σ_j , and $y(t)$ as specified above. Define

$$v(p) = \frac{1}{2} \sum_{j=1}^m (p_j - p_j^*)^2$$

and $M = \min\{p_1^*, p_2^*, \dots, p_m^*\}$. If one has

$$\sum_{j=1}^m p_j^* f_j > \frac{1}{2} \sum_{i,j} \sigma_j(p) \sigma_i(p)$$

then $p(t) \rightarrow p^*$ with probability $1 - v(p)/M$.

Proof: If $p(t)$ is a motion of solution process of (A2), then $v(p)$ is a random function of an Ito process. In the region $Q_M = \{p = |p_i - p_i^*|^2 < M^2\}$, $v(p)$ is a bounded and continuously differentiable process, positive on Q_M and zero only at p^* . We may term $v(p)$ a stochastic Liapunov function. As Kushner has shown, to establish the stability (in probability) result of the theorem, it suffices to show that the time average of $v(p)$ is strictly negative on Q_M . Since $v(p)$ is continuously differentiable on Q_M its time average is computed by means of the differential generator

$$(A3) \quad L = \sum_j f_j(p) \frac{\partial}{\partial p_j} + \frac{1}{2} \sum_{i,j} S_{ij} \frac{\partial^2}{\partial p_i \partial p_j} + \frac{\partial}{\partial t}$$

Here S_{ij} is the i, j th entry in the matrix $S(p) = \sigma(p)\sigma(p)^T$ where $\sigma(p)^T = (\sigma_1(p), \sigma_2(p), \dots, \sigma_m(p))$. Thus,

$$(A4) \quad L v(p) = \sum_j f_j(p) \frac{\partial v(p)}{\partial p_j} + \frac{1}{2} \sum_{i,j} S_{ij}(p) \frac{\partial^2 v(p)}{\partial p_i \partial p_j}$$

⁹ For some indication of this possibility see R. Z. Khas'minskiĭ, "Necessary and Sufficient Conditions for the Asymptotic Stability of Stochastic Linear Systems," *Theory of Probability and Its Applications*, Vol. 12, No. 1, 1967; in particular example 2, p. 146.

¹⁰ See S. J. Turnovsky and E. R. Weintraub, "Stochastic Stability of a General Equilibrium System under Adaptive Expectations," Discussion Paper 123, Dept. of Econ., Univ. of Pennsylvania, June, 1969.

¹¹ H. Bunke, "Stabilität bei Stochastischen Differentialgleichungssystemen," *ZAMM*, Vol. 43, 1963.

¹² Kushner, *op. cit.*, pp. 12-18.

¹³ See footnote 8.

or

$$(A5) \quad Lv(p) = \sum_j f_j(p)(p_j - p_j^*) + \frac{1}{2} \sum S_{ij}(p) I,$$

for I the $m \times m$ identity matrix. Since $\sum_j f_j(p)p_j = 0$ by w , the stability condition $Lv < 0$ reduces to

$$(A6) \quad - \sum_j f_j(p)p_j^* + \frac{1}{2} \sum_{i,j} \sigma_i(p)\sigma_j(p) < 0,$$

which is the assumption of the theorem: i.e., if (A6) is true, $Lv < 0$ and $p(t) \rightarrow p^*(t)$ with probability $1 - v(p)/M$.

FARMER'S RESPONSE TO PRICE IN UNDERDEVELOPED AREAS: THE NICARAGUAN CASE*

By PEDRO BELLI

University of California, Riverside

The behavior of farmers as economic agents has been a subject of some controversy. On the one hand, there are those who believe that farmers, especially farmers in underdeveloped countries, do not respond to price and income incentives [3] [4] [7] [16], and on the other, there are those who maintain that their behavior is in accord with economic theory but that "shift variables" veil the relationship [2] [5] [8] [10].

At first sight it may seem as if the growth of cotton farming in Nicaragua presents a good case in favor of those who do not believe in the rationality of farmers. Cotton has been in the Nicaraguan export ledger at least since the beginning of the nineteenth century, but it was not until 1950 that cotton exports began to acquire some importance. It was around that time that Nicaraguans from all walks of life apparently decided that the road to wealth was paved with cotton. At first only a few intrepid souls ventured into the cultivation of what later came to be known as "white gold," but later, as vast numbers joined, cotton became the leading crop and eventually it displaced coffee as the chief export. Production expanded so fast that by 1966 Nicaragua had become the tenth largest exporter of cotton in the world—a remarkable achievement in view of the country's size and of the initial levels of production.

From 1949 to 1953, the growth in production coincided with a rise in cotton prices, but from 1953 to 1960, cotton prices declined rapidly, from an average of \$32.27 per hundredweight, to an

average of \$24.09 per hundredweight. Exports, which represent over 90 percent of total production, went up from 13,000 metric tons in 1953, to 27,000 in 1960, and finally to 116,000 in 1966. It would seem that during this period Nicaraguan farmers behaved contrary to the tenets of economic theory. Indeed, attempts to estimate the long-run price elasticity of supply were disappointing; neither supply functions without lags, with simple lags, or with more complex distributed lags showed positive results. Estimates of price elasticities in all cases were either abnormally low or negative. These findings were in conflict with those of Krishna [5] and Nerlove [8] [10].

The sparse data available on input prices provided little comfort, for most factor prices either increased or remained constant during the period. Therefore, whatever incentives moved Nicaraguan farmers, it would seem that they were neither price nor income incentives. Yet, as the author hopes to show, it would be erroneous to conclude that these farmers did not behave as economic agents. The rest of this paper is an attempt to show that their behavior is understandable in terms of the theory of the firm. For the purpose, the author presents econometric tests of three hypotheses that provide alternate but not mutually exclusive explanations. The paper ends with a discussion of institutional changes that had profound effects on the development of Nicaraguan agriculture.

As Table 1 shows, there was a phenomenal jump in cotton prices from 1950 to 1951, and even though there was a substantial decline in subsequent years, the average export price from 1950 to 1955 was approximately 50 percent higher than the average of the preceding and following decades. The 1950–55 period, therefore, represents an anomaly in the international cotton market, and it is not at all surprising that it spurred Nicaraguan production to the extent that it did. We may conceive of it as a sort of "big push" period, to use Hirschman's terminology. It is the years from 1955 on that are puzzling, and the remarks that follow attempt to explain the latter period rather than the earlier one.

The assumption that the abnormally high

* This is a partial summary of my doctoral dissertation entitled, "An Inquiry Concerning the Growth of Cotton Farming in Nicaragua," and completed at the Univ. of California, Berkeley, in Aug., 1968, under the supervision of Professors Thomas J. Rothenberg, Daniel L. McFadden, and James N. Bolea. I wish to acknowledge my debt to these gentlemen, whose suggestions and criticisms improved considerably my attempts to answer the questions raised here. I am especially indebted to Professor Rothenberg, who provided so many ideas that any merits that this paper may have should be credited to him. Needless to say, whatever errors remain because I was either too stubborn or too obtuse to remove are entirely my responsibility. I also wish to acknowledge the generous financial aid received from the Institute of International Studies and from the Center for Latin American Studies, both of the Univ. of California, Berkeley.

TABLE 1
NICARAGUAN COTTON: QUANTITY EXPORTED AND
AVERAGE VALUE OF EXPORTS*

Year	Exports (Metric Tons)	Average Value† (Dollars per cwt.)
1910	63	18.63
1920	77	15.91
1930	147	31.53
1940	1,088	20.00
1950	3,307	29.09
1951	4,358	59.09
1952	9,530	34.54
1953	12,783	32.27
1954	23,196	35.45
1955	43,972	33.63
1960	27,390	24.09
1961	32,515	25.45
1962	55,660	25.45
1963	73,125	24.54
1964	93,484	24.54
1965	121,631	24.09
1966	116,000	21.36

* Calculated from: Nicaragua, Oficina del Recaudador General de Aduanas, *Memoria del Recaudador General de Aduanas* (Managua: Oficina del Recaudador General de Aduanas), various issues, and from U. S. Department of Commerce, Bureau of the Census, *Statistical Abstract of the United States* (Washington: U. S. Government Printing Office), various issues.

† Deflated values. From 1904 to 1950, all values are deflated by the U. S. Bureau of Labor Statistics Index of Wholesale Prices of all Commodities. From 1950 to 1966 by a consumer price index of Nicaragua. The year 1958 is the base year for both indices.

prices of the 1950-55 period provided farmers with ample protection is common to all hypotheses tested here. During this period numerous people with little knowledge of agriculture tried their hands at growing cotton, and very likely they made numerous mistakes. But the price umbrella prevented total disaster. It is reasonable to suppose that during this initial period the neophytes learned something about the cultivation of the crop, and that this knowledge served them well later on.

Returns to Scale

The first test concerned the hypothesis that among the many things that farmers learned from 1950 to 1955 was that average costs decrease as the size of the farm increases. Under these conditions, in order to maintain or augment a given profit margin, a rational farmer would have been inclined to cultivate increasingly larger farms.

A priori, then, we would have expected to find a tendency towards concentration in the industry. Nevertheless, as Table 3 shows, such a movement is virtually absent. In spite of the growth in the number of cotton farmers, the percentage distribution of farms by size remained impressively constant, suggesting that cotton production is not subject to increasing returns to scale. Econometric attempts to measure returns to scale confirm this conclusion.

A cross-section sample of 311 farms showing expenditures on various items during the crop year 1964-65 provided the basis for estimating the degree of returns to scale. Because space limitations do not permit a detailed explanation of the estimating techniques, only a broad outline will be presented here. Using a Cobb-Douglas function as an empirically useful approximation to the production function, the author attempted to measure the degree of homogeneity of the latter. Estimates of the scale parameter indicated that increasing returns to scale do not exist but that constant returns seem to be more likely.

Because there are reasons to believe that there is a simultaneous equation bias in estimating production functions [6] [9], the author used three estimating procedures, of which the last two presumably yield consistent estimates of the scale parameter. These methods were: least squares on the production function, Klein's relative shares

TABLE 2
NICARAGUA: INDICES OF ESTIMATED PRICES OF
INPUTS USED IN COTTON PRODUCTION
(1964-65=100)*

Crop Year	Land Rental†	Labor†	Insecti- cide†	Fertil- izers†
1951-52	—	—	—	204
1952-53	—	—	—	107
1953-54	—	53	—	100
1954-55	—	37	—	176
1955-56	—	57	—	104
1956-57	—	47	—	104
1957-58	52	46	—	85
1958-59	53	55	—	83
1959-60	110	74	103	97
1960-61	100	98	103	92
1961-62	67	97	106	114
1962-63	72	99	116	106
1963-64	102	106	110	104
1964-65	100	100	100	100
1965-66	113	93	97	105

* Calculated from own field research.

† A dash indicates that the data were not available.

TABLE 3
NICARAGUA: PERCENTAGE DISTRIBUTION OF FARMS BY SIZE*

Farm Size† (Hectares)	Crop Year							
	1951-52‡	1960-61	1961-62	1962-63	1963-64	1964-65	1965-66	1966-67
0- 35	69.81	76.42	79.23	79.28	80.09	76.73	76.38	72.42
35- 70	13.03	10.92	8.72	8.81	9.04	10.79	10.06	11.22
70-140	13.56	7.99	7.75	7.19	6.57	7.29	7.95	9.35
140-350	3.6	4.22	3.88	4.20	3.77	4.48	4.72	5.86
Over 350	0.0	0.45	0.42	0.52	0.52	0.72	0.89	1.14
Number of farms	1,305	2,015	2,889	3,475	4,426	4,469	5,081	4,384

* Calculated from: Nicaragua, Dirección General de Estadística y Censos, *Boletín de Estadística*, III Época, No. 4, pp. 49-54, Nicaragua, Ministerio de Agricultura y Ganadería, División de Estudios Económicos y Agropecuarios, and Nicaragua, Comisión Nacional del Algodón.

† These intervals are open on the right.

‡ This column is a crude estimate based on census data. It is not strictly comparable to the other columns.

method [9], and indirect least squares, using direct least squares on the supply function.

The results were very similar with all three methods. The sum of the production function coefficients was always close to unity, indicating constant rather than increasing returns.¹

It is interesting to note that of the three values estimated for the short-run price elasticity, two were well within the range of those estimated by Nerlove [10], Krishna [5], and Falcon [2].² The one obtained by Klein's method appears to be much too high. Nevertheless, all of them tend to

contradict the hypothesis that farmers do not respond to price. Table 4 is a summary of the results.

Technical Change

The second hypothesis tested concerned the possibility that costs of production declined through time because of technical change. Unfortunately, the scarcity of data did not permit a good test and the results were inconclusive.

Because the author had at his disposal a time series of cross-section sample, the analytic technique was slightly different from the usual one encountered in econometric work, and slightly more general.

Specifically, beginning with the functions,

$$(1) \ln Y_{ij} = \delta_i + \beta_j + D + \sum_k \alpha_k \ln X_{ijk} + \gamma L_{ij} + \epsilon_{ij}$$

TABLE 4
ESTIMATES OF THE DEGREE OF HOMOGENEITY OF THE PRODUCTION FUNCTION
AND OF THE SHORT-RUN PRICE ELASTICITY OF SUPPLY

	Method of Estimation		
	Least Squares on Production Function	Indirect Least Squares	Klein's
Scale parameter standard error	1.058 (0.0015)	1.043 (Not Available)	1.062 (Not Available)
Price elasticity of supply	0.66	0.42	1.37

¹ Estimates of individual production function coefficients, however, differed substantially from method to method.

² These values were: 0.34, Nerlove (U.S.); 0.72 and 0.59, Krishna (Punjab); and 0.41, Falcon (Punjab).

TABLE 5
ESTIMATES OF THE SHIFT PARAMETER, β_j

	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965
Value*.....	0	-2.80	5.82	7.90	1.77	8.24	2.62	9.03	11.62	11.58	13.83	7.23
Standard error.		5.12	5.34	6.79	6.08	6.70	6.65	6.79	6.96	6.42	6.14	6.95
Value†.....			0	0.16	-0.28	0.79	-0.17	0.11	0.33	0.25	0.26	0.16
Standard error.				0.20	0.21	0.23	0.23	0.23	0.26	0.24	0.21	0.25

* Estimates obtained from equation (2).

† Estimates obtained from equation (1). The same number of years do not appear in both equations because one farm, of the two available for those years, did not use fertilizer during 1954 and 1955. Because the Cobb-Douglas version of the production function is not pliable enough to handle problems like this, the author chose to eliminate these years from the analysis.

$$(2) \quad Z_{ij} = \delta_i + \beta_j + D' + \sum_k \lambda_k (X_{ijk}/L_{ij}) + \mu L_{ij} + u_{ij}$$

where λ denotes output of cotton, δ and β denote two shift parameters, D , D' , α_k , λ_k , γ and μ stand for production function coefficients, X_k for the amount of the k th input, L for the area cultivated, Z for the amount of output of cotton per hectare (henceforth, yield), ϵ and u for random errors, and where i and j indicate the i th farmer and the j th year, respectively, the test consisted in ascertaining whether the hypothesis, $\beta_1 = \beta_2 = \dots = \beta_t = 0$ could be rejected at the 5 percent level. As Tables 6 and 7 indicate, the null hypothesis could not be rejected at the specified level. On the other hand, as Table 5 shows, estimates of each β_j increased through time and from this table alone it is difficult to resist the temptation of rejecting the null hypothesis on a purely intuitive level. Perhaps

the large standard error associated with each β_j simply mean that the measurements are imprecise, not that the actual values of the shift parameters are zero. In view of size of the sample (observations on six farms over periods ranging from five to twelve years, for a total of 36 observations), the author is inclined to think that there was some technical change, but that it was not strong enough to show up in the data.

Learning by Doing

In his article on learning by doing, Kenneth Arrow [1] mentioned that two generalizations from psychological learning theories emerge with particular interest for economists. One, learning appears as a result of repetition and two, "... learning associated with repetition ... is subject to sharply diminishing returns. ..." These findings

TABLE 6
ANALYSIS OF COVARIANCE, TWO-WAY CLASSIFICATION, EQUATION (1)

Source of Variation	R^2	Adjusted Sum of Squares for Dependent Variable	D.F.	Mean Square	F Ratio
(1)	(2)	(2)	(3)	(4)	(5)
1. Between farms.....		0.841715	5	0.16834	5.97
2. Between years.....		0.346793	9	0.03853	1.37
3. Residual when model contains farm effects only.....	.9292	0.805306	20		
4. Residual when model contains year effects only.....	.8988	1.152099	16		
5. Residual when model contains both farm and year effects...	.9727	0.310384	11	0.0281	

TABLE 7
ANALYSIS OF COVARIANCE, TWO-WAY CLASSIFICATION, EQUATION (2)

Source of Variation	R^2	Adjusted Sum of Squares for Dependent Variable	D.F.	Mean Square	F Ratio
(1)	(2)	(2)	(3)	(4)	(5)
1. Between farms.....		625.65	5	125.13	5.81
2. Between years.....		384.29	11	34.94	1.63
3. Residual when model contains farm effects only.....	.8186	729.24	27		
4. Residual when model contains year effects only.....	.7587	969.90	21		
5. Residual when model contains both farm and year effects...	.9144	344.25	16	21.52	

suggest that profit maximization is a seldom-attained goal. Even at constant factor and output prices, an entrepreneur learning a new business discovers the optimum input mix only after a few trials. When relative prices change, the search for the profit maximizing input mix begins anew. In an environment with unstable prices this quest becomes a continuous pursuit of elusive goals. This section is an attempt to bring these two generalizations into the usual profit maximizing scheme in an effort to explain the growth of the Nicaraguan cotton industry as a case of learning by doing.

Three basic assumptions underlie the analysis. One, that entrepreneurs undertaking a new venture only by chance attain the profit maximizing input mix. Two, that with practice (learning by doing) they attain it. And three, that input levels also depend on factor prices. In symbols:

$$(3.a) \quad \lim_{t \rightarrow \infty} X_{it} = \bar{X}_i$$

$$(3.b) \quad \lim_{t \rightarrow \infty} \frac{\partial X_{it}}{\partial t} = 0$$

$$(4) \quad X_{it} = X_i(P_0, P_1, \dots, P_n; t)$$

where X_{it} stands for the actual observed level of the i^{th} input in the t^{th} trial, \bar{X}_i for the profit maximizing level of X_i , P_i for the price of X_i , and n for some finite number that indicates the number of trials necessary for "learning" to occur. Equation (3.a) reflects the generalization that learning appears as a result of repetition and equation (3.b) that learning is subject to diminishing returns.

As an empirical approximation incorporating all three conditions, the author used the following equation:

$$(5) \quad \ln X_{ijt} = C + \delta_j + a \frac{P_t}{P_0} + bi \frac{1}{t} + \epsilon_{jt}$$

Where C and a are two constants, b stands for a "learning coefficient," δ for a shift parameter introduced in order to compensate for possible differences in speeds of learning among farmers, P_0 for the price of cotton, and ϵ for random error. t denotes the trial number, j denotes the farmer, and i denotes the input. The meaning of the other symbols remains the same as before.

Ideally, t should have been equal to unity the first time that a particular farmer grew cotton, and it should have increased by one unit every time that he reaped a new crop. Such data were not available and hence here $t=1$ for the first datum available on a particular farmer, not necessarily the first trial.

Where learning was an influential factor, the estimate of b should have been different from zero. Moreover, given the rapid expansion from 1952 to 1957 and from 1960 to 1965, it is reasonable to suppose that most of the essential learning, if any, occurred during the first five years, and that after ten years, at constant factor prices, farmers would have mastered the techniques of production. Hence, as a test of the learning-by-doing hypothesis the author required not only that t be an influential factor, but also that the estimated value of b should suggest that the amount

TABLE 8
ESTIMATES OF THE COEFFICIENTS OF THE LEARNING BY
DOING DERIVED DEMAND FOR INPUTS

Dependent Variable	Coefficient		
	Constant	a	b
Land.....	5.305693 (0.399616)	6.751101 (10.954935)	-0.826932 (0.476186)
Fertilizer per unit of area.....	5.919489 (0.601217)	-1.024522 (1.252299)	-2.971578 (1.098425)
Insecticide per unit of area...	6.852612 (0.848334)	0.108754 (0.387599)	-3.318659 (1.010882)
Labor per unit of area.....	6.661010 (0.421385)	1.214550 (1.170558)	-1.449347 (0.764960)
<i>D.F.: 27</i>			

used of any particular input at $t=10$ should be no more than 10 percent away from the profit maximizing input level. In equation (5) this would imply that the estimate of b should be around one: negative if farmers used too little of the input in the beginning; positive if they used too much. Table 8 is a summary of the results.

The estimates of b were, in all cases, different from zero. The hypothesis that $b=-1$ could be rejected only in the case for insecticides, at the 5 percent level. Hence, we cannot rule out learning by doing as a contributing factor in the growth of the industry. Nevertheless, the estimates of b for fertilizers and insecticides were large, suggesting that the learning process was slower than was reasonable to expect.

In summary, econometric tests suggest, one, that cotton farming is subject to constant returns to scale. Two, that technical change may have played a part in the expansion of the sector, but that if it did, it was not too important, and, finally, that learning by doing occurred, but that it was slower than expected. In short, neither the conditions under which cotton is produced, nor technical advances affecting cotton production, nor learning by farmers explain (alone or in conjunction) the fast development of cotton farming. Had the industry grown more slowly, the results obtained might have been more convincing. Had the results been stronger (sharply increasing returns, or faster learning, for example) the explanation based upon them may have been more convincing. But when all is said and done, one still has the feeling that there is a missing ele-

ment, that something important has been left out.

The missing element—the catalyst that brought the development mentioned above to fruition and that opened the farm doors to those who theretofore had had the ability but not the monetary means to till the land—was the liberalization of the agricultural credit laws.

The Expansion of Agricultural Credit

For reasons that are not clear at this point, in 1940 the Nicaraguan government undertook the complete reorganization of the laws governing the Banco Nacional de Nicaragua and all the commercial banks in an attempt to pay attention to the role that credit could play in the economic development of the country. Up until that time all Nicaraguan banks, including the Banco Nacional, were prohibited by law from extending credit to any one farmer beyond a certain limit that in fact was enough to cultivate only around 70 hectares [11] [12] [13] [14] [15]. Because the majority of cotton farmers, then and now, work primarily on credit, this ceiling put an effective limit on the size of the farm that most farmers could cultivate. Banking laws, moreover, prohibited loans to tenant growers. Successive modifications of banking laws in 1941, 1949, and 1952 increased the credit limit enough to cover the costs of cultivating more than 170 hectares. In the process, the government finally recognized that tenant growers also needed money and modified the laws accordingly. When the price of cotton jumped from \$29.09 to \$59.09, many Nicaraguans

turned to agriculture. Fortunately, wealth was not a prerequisite for cultivating the land, for the Banco Nacional stood ready to finance future cotton growers, with the future crop as the only security. Agricultural credit rose phenomenally and the Banco began to grant more liberal loans to farmers, as Table 9 shows.

The laws described above remained in effect until 1959. After bad weather caused crop failures and low cotton prices began to take a toll of high-cost farmers and to strain the reserves of the Banco, the latter put into effect a new credit system designed to winnow the inefficient entrepreneur. Thus, only cotton growers whose average yield over the previous three years exceeded 30 hundredweight per hectare could obtain loans from the Banco. Moreover, credit limits were eased; any farmer could borrow up to 70 percent of the estimated cost of cultivating as much land as he saw fit [11]. Table 3 reflects the results of these reforms.

Thus, it was a combination of changes in the country's credit policies, the presence of constant returns to scale, and the results of learning by doing that spurred the growth of the industry. From 1951-52 to 1954-55 (crop years) the price of cotton remained stable. Under constant returns and in the absence of institutional barriers to entry, we would have expected the industry to expand. It did. Output went up from 8,700

metric tons to 44,400 (there is no data available on the number of firms in the industry during these years).

From 1955-56 to 1959-60 cotton prices fluctuated slightly, but the general trend was downwards. The industry's output contracted; output went down from 34,400 to 28,800 metric tons. After 1960 prices began to go up and the reforms of the banking laws to sift the low-yield farmers. It was during these latter years of slightly rising prices that the industry expanded most rapidly. This is what we would have expected in an industry with constant returns and in which new entrepreneurs were beginning to master the techniques of production, especially if it is granted that by 1960 farmers had learned enough about cotton farming to be able to reap the full benefits of constant costs.

In conclusion, in spite of appearances to the contrary, the growth of cotton farming in Nicaragua is a case in favor of those who believe that farmers do respond to price and income incentives. It seems that the output of cotton increased in spite of falling prices because there were changes in agricultural credit policies that opened the farm door to those who theretofore had had the desire but not the money to become cotton farmers. Second, constant returns to scale—with its concomitant absence of cost barriers to entry—permitted expansion of the industry during periods of stable prices. Thus, output grew, not only because old firms in the industry were able to increase their scale of operations without incurring higher average costs, but more importantly, because new firms coming into the industry could produce, at least theoretically, at the same average costs as those already in it. Third, the data lend credence to the hypothesis that average costs declined as a consequence of learning by doing. Finally, there are some indications, however weak, that technical change may also have contributed to the reduction of average costs. As Nerlove [10], Krishna [5], and Falcon [2] have shown before, the behavior of farmers was quite consistent with the tenets of economic theory but "shift variables" hid the true relationship between price and output.

REFERENCES

1. KENNETH ARROW, "The Economic Implications of Learning by Doing," *Rev. of Econ. Studies*, June, 1962, pp. 155-73.
2. WALTER P. FALCON, "Farmer Response to Price in a Subsistence Economy: The Case of West Pakistan," *A.E.R.*, May, 1964, pp. 580-91.
3. JOSEPH GRUNWALD, "The Structuralist School on

TABLE 9
AGRICULTURAL LOANS OF THE BANCO NACIONAL
DE NICARAGUA, SELECTED YEARS*

Crop Year	Loans to Cotton Farmers		
	Agricultural Loans (Cordobas ×1000)	Total (Cordobas ×1000)	Average per Hectare (Cordobas/Hectare)
1944-45	5,109	762	423
1948-49	13,365	1,005	421
1950-51	21,068	11,386	592
1952-53	63,485	22,560	998
1954-55	140,865	94,844	1,425
1956-57	147,428	89,604	1,473
1958-59†	141,680	84,282	1,448
1960-61†	107,936	63,691	1,619
1962-63†	160,785	107,972	1,702
1964-65†	248,236	180,416	1,840

* Calculated from Nicaragua, Banco Nacional de Nicaragua, *Revista Trimestral del Banco Nacional de Nicaragua*, various issues.

† These years include loans by commercial banks as well.

- Price Stability and Development," in A. O. Hirschman, ed., *Latin American Issues* (Twentieth Century Fund, 1961).
4. D. R. KHATKHATE, "Some Notes on the Real Effects on Foreign Surplus Disposal in Underdeveloped Areas," *Q.J.E.*, May, 1962, pp. 186-96.
 5. RAJ KRISHNA, "Farm Supply Response in India-Pakistan," *Econ. J.*, Sept., 1963, pp. 477-87.
 6. J. MARSCHAK AND W. H. ANDREWS, JR., "Random Simultaneous Equations and the Theory of Production," *Econometrica*, 1944, pp. 143-205.
 7. KUSUM NAIR, *Blossoms in the Dust* (London: Duckworth & Co., Ltd., 1961).
 8. MARC NERLOVE, "Estimates of the Elasticities of Supply of Selected Agricultural Commodities," *J. of Farm Econ.*, May, 1956, pp. 496-509.
 9. ———, *Estimation and Identification of Cobb-Douglas Production Functions* (Rand McNally & Co., 1965).
 10. ———, *The Dynamics of Supply: Farmers' Response to Price* (Johns Hopkins Press, 1958).
 11. NICARAGUA, BANCO NACIONAL DE NICARAGUA, *El Banco Nacional de Nicaragua y Sus Programas de Desarrollo* (Managua: Litografía San José, 1966).
 12. ———, *Manual de Informaciones Estadísticas*, Managua, n.d. (mimeographed).
 13. ———, *Revista Trimestral del Banco Nacional de Nicaragua* (Managua: Banco Nacional), various issues.
 14. ———, *Estudio de la Economía del Algodón en Nicaragua*, Managua, n.d. (mimeographed).
 15. ———, *Leyes Bancarias y Monetarias: República de Nicaragua* (Managua: Talleres Tipográficos Heuberger, 1951).
 16. R. O. OLSON, "Discussion: Impact and Implications of Foreign Surplus Disposal in Underdeveloped Economies," *J. of Farm Econ.*, Dec., 1960, pp. 1042-45.

EQUILIBRIUM GROWTH AND CAPITAL MOVEMENTS BETWEEN OPEN ECONOMIES*

By MYRON G. MYERS
Rutgers University

In his well-known article written four decades ago, J. H. Williams [5] voiced a criticism of international trade theory that remains largely valid today: "the . . . relation of international trade to the development of new resources and productive forces is a more significant part of the explanation of the present status of nations, of incomes, prices, well-being, than is the cross-section value analysis of the classical economists, with its assumption of a given quanta of productive factors, already existent and employed; . . . that the international movements of productive factors has significance relative to comparative prices, incomes, positions of nations, at least equal to that of the trade in goods, . . ."

Most familiar analytical trade models of the two-country variety are either static, or when dynamic, posit balanced current accounts over time.¹ Few treatments of dynamic issues assign a role to international investment and the concomitant service flows to which it gives rise as these factors relate to the time paths of incomes, prices and levels of international trade.²

The distinction between growth models of a closed and open nature when foreign investment is free to occur may be seen by considering the savings-investment decision. In a simple closed neoclassical growth model with only one asset in existence—claims over real capital—a decision not to consume automatically implies a decision to invest and increase the real domestic stock of capital. In contrast, an open economy, with only a single asset in a world where capital is free to move, has two potential outlets for domestic savings—domestic or foreign investment. In equilibrium growth, the proportionate growth rate of output (\dot{X}/X) equals the natural rate of growth (n) and the rate at which capital is accumulated I/K (where I is investment and K the

stock of capital)

$$\frac{\dot{X}}{X} = n = \frac{I}{K} = \frac{I}{X} \frac{X}{K}.$$

The equilibrium output-capital ratio X/K equals the natural rate of growth divided by the investment-output ratio.

$$\frac{X}{K} = \frac{n}{I/X}$$

In an open economy it no longer necessarily follows that the portion of output saved, S/X , coincides with the fraction invested domestically, I/X . The steady-state output-capital ratio and hence the long-run equilibrium in general can be expected to undergo considerable change when savings performed domestically are transmitted abroad or when the economy receives a capital inflow from abroad. In addition, national income is now distinct from domestic product due to the debt-cum-service factor arising from past levels and directions of international investment. To the extent that such a flow of income alters world expenditure patterns, both the terms of trade and the time path of world capital accumulation may be affected by international investment and by the resultant future flow of interest and dividends.

In order to establish the necessary analytical framework to deal with the problems posed above, a dynamic variant of the familiar two-country (each with single sectors), two-commodity, two-factor international trade model is delineated.

I. The Model

The world economy is comprised of two countries, denoted by the subscripts 1 and 2, with the output of the first (Agraria) consisting wholly of the homogeneous consumer good X , while the second (Mancunia) is specialized in the production of the infinitely long-lived homogeneous producers good identical with the factor capital, Z . The form of the functions governing production in both economies is specified as Cobb-Douglas with the parameters a and b , the respective output elasticities of capital or capital's (constant) share of output when factors are paid

* This paper is a summary of a portion of my Ph.D. dissertation submitted to the faculty of Brown Univ. Appreciation is extended to G. H. Borts, R. Sato, J. Hanson, H. Ryder, Jr., E. R. Weintraub and P. Asch for their many helpful suggestions and comments.

¹ See for example Bardhan [1] and Oniki and Uzawa [4].

² For a previous attempt using a world composed of two countries that produce the same all-inclusive output with the same technology, see Hamada [2].

their marginal products.

$$(1) \quad X = F_1(K_1, L_1) = K_1^a L_1^{1-a}$$

$$(1') \quad Z = F_2(K_2, L_2) = K_2^b L_2^{1-b}$$

The subscripted variables K_i and L_i denote the capital and labor inputs of which positive levels are required to produce output.

The effective population, assumed identical to the labor force in each country, grows at a demographically given constant proportionate rate n_i ($i=1, 2$) and is assumed internationally immobile.

$$(2) \quad L_1(t) = L_1(0) \exp(n_1 t)$$

$$(2') \quad L_2(t) = L_2(0) \exp(n_2 t)$$

The following set of six equations serves to define the source and ultimate use of the two world commodities.

The domestic product of Agraria (X) may either be consumed at home (X_c) or exported to Mancunia (X_a) where it is consumed, while Mancunia's annual product (Z) is either retained at home (Z_c) or exported to Agraria (Z_a).

$$(3) \quad X = X_c + X_a$$

$$(3') \quad Z = Z_c + Z_a$$

The current flow of Mancunia's newly produced output is used to augment the existing indestructible stock of capital in each country.

$$(4) \quad \dot{K}_1 = Z_a$$

$$(4') \quad \dot{K}_2 = Z_a$$

A salient feature of a system composed of open economies is the potential divergence between national incomes and domestic products. When international investment takes place, the current value of national income of an economy will exceed the current value of domestic product if residents have on balance invested more abroad than foreigners have invested domestically. More precisely, national income exceeds domestic product by the amount of income composed of interest and dividends on the net total of previously lent capital employed abroad. Conversely, a debtor state is obliged to transfer abroad a portion of its domestic product representing interest and dividends to owners or lenders of domestically employed capital. Define $P(t)$ as the relative price of X in terms of Z , $r(t)$ as the interest rate prevailing in the world and $D(t)$ —measured in units of Z —as the stock of international indebtedness, all at time t . National income, measured in terms of the numeraire Z , is domestic product plus income from the ownership of claims on capital employed abroad or minus the income accruing to foreign

holders of claims over domestic capital if the economy is a net creditor or debtor, respectively.

$$(5) \quad PX - rD = S_1 + PX_c$$

$$(5') \quad Z + rD = S_2 + PX_a$$

Equations (5) and (5') are the usual social accounting identities on the disposition of national income. Domestic product—adjusted by either receipts from abroad or payments to foreigners—is either saved (S_i) or consumed where consumption is limited to Agraria's current product. By convention, the stock of international debt, $D(t)$, is positive when Agraria has been the net recipient over time of investment from Mancunia and negative when Mancunia is the world debtor. For analytical simplicity, $D(t)$ represents an equity instrument so that the flow of service payments at time t is the stock of equity debt $D(t)$ times the prevailing world interest rate $r(t)$.

Savings is assumed to depend functionally on the level of national income with a constant proportion s_i of national income saved.

$$(6) \quad S_1 = s_1(PX - rD) \quad 0 < s_i < 1 \quad (i = 1, 2)$$

$$(6') \quad S_2 = s_2(Z + rD)$$

The next set of relationships embodies the assumption of an unfettered international capital market where capital flows between countries so as to instantaneously equate marginal returns to capital, $r(t)$.

$$(7) \quad r(t) = \frac{\partial F_1}{\partial K_1} \cdot P(t)$$

$$(7') \quad r(t) = \frac{\partial F_2}{\partial K_2}$$

To close the model, the dynamic link between the historic behavior of borrowing and lending and its effect on the present must be specified. An import of capital today is ultimately followed by a reverse flow of service payments (interest and dividends) in the future. The service flows appear as a debit in the current account of the debtor country and as a corresponding credit in the current account of the creditor nation. The change in a country's creditor-debtor position depends on the magnitude of its current account surplus or deficit. Equation (8) defines the change in the stock of world indebtedness as the negative of Agraria's current account and is determined by the current levels of exports and imports as well as the size of the existing stock of debt. If $\dot{D}(t)$ is positive when $D(t)$ is positive, Agraria's imports plus remittances to residents of Mancunia exceed its exports and there is an increase in foreign liabilities. When $\dot{D}(t)$ is positive and $D(t)$ negative, Agraria's imports exceed the com-

bined value of exports and service receipts from abroad and there is a decrease in domestically owned claims on foreigners. Appropriate results are obtained for the instance when $\dot{D}(t)$ is negative in which case Mancunia's liabilities to residents of Agraria increase or foreign assets decrease depending on whether $D(t)$ is negative or positive, respectively.

$$(8) \quad \dot{D} = Z_s + rD - PX_s$$

In summary, a capital flow from Mancunia to Agraria occurs when Agraria has a deficit on current account (surplus on capital account) with the transfer of resources in excess of earning matched by an opposite flow of titles to ownership over productive resources.

The model is complete and consists of fifteen equations in fifteen unknowns. Equations (3), (4), (5) and (8) may be combined to yield $\dot{D} = K_1 - S_1$ and similarly, using (3'), (4'), (5') and (8), $\dot{D} = S_2 - K_2$ —when investment exceeds savings in Agraria, savings exceeds investment in Mancunia and an import of capital to Agraria coincides with an export of capital from Mancunia.

II. The Short-run Equilibrium

Define $K(t)$ and $L(t)$ as the momentary world endowments (sum of the existing stocks in each country) of capital and labor, respectively. In addition, let:

$$k = K/L$$

$$d = D/L$$

$$k_i = \frac{K_i}{L_i} \quad (i = 1, 2)$$

$$l_i = \frac{L_i}{L} \quad (i = 1, 2)$$

It will prove possible to solve explicitly for the values of the required capital stocks in each economy in terms of the momentary levels of the world stocks of capital $K(t)$ and indebtedness $D(t)$. The allocation of the world endowment of capital must be consistent with the requirements that marginal returns to capital are everywhere equal. In order to solve for these short-run terms, and then to incorporate them into the dynamic behavior relationship as stipulated in the model, it is first necessary to establish certain other relationships.

The terms of trade or relative price ratio $P(t)$ may be found by equating the offer of the consumption good by Agraria to the demand for the consumption good by Mancunia as determined by (5), (6) and (5'), (6'), respectively. Alternatively, the balance-of-payments constraint written as world savings equals world investment can

be used to give the terms of trade.

$$(9) \quad P = \frac{(1 - s_2)Z + (s_1 - s_2)rD}{s_1X}$$

Note that service flows, for example, from Agraria to Mancunia ($D > 0$) have the effect of turning the terms of trade in favor of Agraria provided $s_1 > s_2$, and in favor of Mancunia otherwise. Only if the disposition of a unit's worth of income between savings versus consumption is internationally the same ($s_1 = s_2$) will past levels of foreign investment have no effect on the current level of the terms of trade.³

The value of Agraria's exports in terms of Mancunia's output and the magnitude of service payments can also be computed by combining (5') and (6').

$$(10) \quad PX_s = (1 - s_2)(Z + rD)$$

Finally, the instantaneous equality of marginal returns on investment in conjunction with the terms of trade determines how the momentary world stock of capital must be apportioned between the two economies. Combining (7), (7') and (9) yields:

$$(11) \quad K_1 = \frac{a(1 - s_2)K - ab(s_2 - s_1)D}{a(1 - s_2) + bs_1}$$

$$(12) \quad K_2 = \frac{bs_1K + ab(s_2 - s_1)D}{a(1 - s_2) + bs_1}$$

It follows from (11) and (12) that the momentary values of K and D determine the allocation of resources between economies and relate the time paths of K and D to the time paths of each economy's capital stock. Also note that identical savings propensities imply that changes in the sign and magnitude of the world stock of indebtedness have no effect on resource allocation and thus on the process of world capital accumulation. The reason for this is the same as the discussion above on the effect of service flows on the terms of trade.

III. Dynamics

The discussion will be limited to the case where labor force growth rates coincide internationally, and the resulting equilibrium motion is such that all variables grow at this coincident rate ($n_1 = n_2 = n$), or else are constant. In particular, the asymptotic convergence of the ratios d and k imply from the short run that capital stocks K_i grow at the rate n . It immediately follows that physical outputs also grow at the rate n while the stock of international indebtedness as a fraction of domestic physical outputs stabilizes. Note

³ This point is raised by Johnson in [3].

also that the convergence of capital-labor (and output-capital) ratios imply that the terms of trade and world interest rate are asymptotically constant so that service flows are annually constant fractions of the creditor and debtors domestic product. Hence, national incomes also grow at the rate n .

The process of world capital accumulation and the concomitant behavior of international investment can be described by a system of two differential equations in a like number of unknowns, obtained by integrating the dynamic structural and behavioral equations of the model with the short-run relationships (9)–(12). The time derivative of the per capita stock of world capital, \dot{k} , depends ultimately on the allocation of labor and capital to Mancunia. Symbolically,

$$\dot{k} = \frac{\dot{K}}{L} - nk = \frac{Z}{L} - nk = l_2 k_2^b - nk$$

From (12), k_2 is related to the levels of k and d resulting in \dot{k} as a function, G , of k and d .

$$(13) \quad \dot{k} = G(k, d) = l_2 \left[\frac{C_1}{l_2} k + \frac{C_2}{l_2} d \right]^b - nk$$

where

$$C_1 = \frac{bs_1}{a(1-s_2) + bs_1} \quad C_2 = \frac{ab(s_2 - s_1)}{a(1-s_2) + bs_1}$$

Similarly, the time derivative of per capita world indebtedness depends, from (8), on the current account balance the components of which are available from the short-run expressions (9)–(12). Alternatively, the same results can be obtained from the identity $\dot{D} = K_1 - S_1 = S_2 - K_2$ and substitution of (6) and (11) or (6') and (12). The results of these considerations determine the equation, F , describing \dot{d} in terms of k and d .

$$(14) \quad \dot{d} = F(k, d) = l_2 C_3 \left[\frac{C_1}{l_2} k + \frac{C_2}{l_2} d \right]^{b-1} + d(s_2 C_4 - n)$$

where

$$(15) \quad r = r(k, d) = b \left[\frac{C_1}{l_2} k + \frac{C_2}{l_2} d \right]^{b-1}$$

and

$$C_3 = \frac{(1-s_2)(as_2 - bs_1)}{a(1-s_2) + bs_1 + ab(s_2 - s_1)}$$

$$C_4 = \frac{a(1-s_2) + bs_1}{a(1-s_2) + bs_1 + ab(s_2 - s_1)}$$

Equations (13), (14) and (15) form the system describing the interdependent motion over time of k and d .

IV. Results

Analysis of the motion of k and d over time, as described by (13), (14) and (15), establishes conditions that guarantee the attainment of a steady-state equilibrium and the characteristics of the equilibrium as influenced by each country's savings behavior and technology. Illustrative phase diagrams are presented below for selected cases of the possible parametric arrays.

In general, the long-run debtor-credit relationship depends on international differences in the ratio of the share of output paid to capital (capital coefficient) to savings rate. As the relative difference in the ratios of capital coefficient to savings rate enlarges, international investment assumes a position of greater importance in the world economy with capital flowing from the country with the higher savings rate and lower capital coefficient to the country with the lower savings rate and higher capital coefficient. As a limiting case of this proposition it can be established that when the ratios of capital coefficient to savings rate coincide internationally, the steady-state solution is characterized by balanced trade; that is, international indebtedness grows at a slower rate than domestic outputs and asymptotically approaches zero while each country's savings ratio coincides with the ratio of output invested domestically.

The case of differing technologies and identical savings rates is depicted in Figure 1 where, without loss of generality, b exceeds a . The sign of the equilibrium stock of world per capita debt (d) is

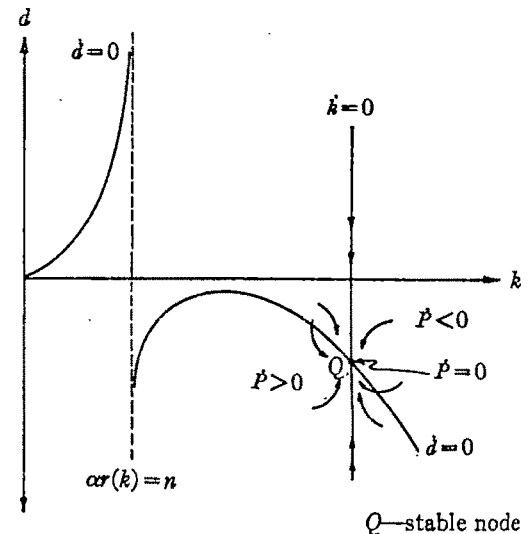


FIGURE 1

negative indicating that Agraria annually lends Mancunia a constant fraction of its national income and annually receives from Mancunia as interest a sum amounting to a constant fraction of its national income.⁴ Superimposed on the phase diagram is the direction taken by the terms of trade as the steady-state is reached. Of course, in the steady-state configuration there is no movement in the terms of trade ($\dot{P}/P=0$). The equilibrium levels of both countries, output-capital ratios can be shown to satisfy the inequality

$$\frac{Z}{K_2} < \frac{\pi}{s} < \frac{PX}{K_1}$$

substantiating the introductory discussion on the effects capital movements have on the receiving and transmitting economies.

Figure 2 is an illustrative phase diagram for the instance when both technology and savings behavior differ internationally. Numerical simulation was required in order to establish that of the two potential equilibria; E_1 is locally stable while E_2 is a saddle point.⁵ At E_1 , the equilibrium d is positive indicating that Mancunia annually transmits capital to Agraria and annually receives service payments from Agraria. Again, the asymptotic convergence of k and d along with the convergence of the common marginal return on investment and relative prices assures that the capital importer annually borrows and annually

⁴ Stability of the steady-state equilibrium is guaranteed by restricting capital coefficients and savings rates to the unit interval.

⁵ Phrased in terms of the equilibrium world rental on capital, r^* , the steady-state is locally stable if $C_1 r^* < \pi$ and $s_2 C_1 r^* < \pi$ provided $0 < b < 1$. When these conditions are met, it is easily established that the sign of the equilibrium stock of debt and hence the direction of capital movements is determined by the sign of $[a/s_1 - b/s_2]$.

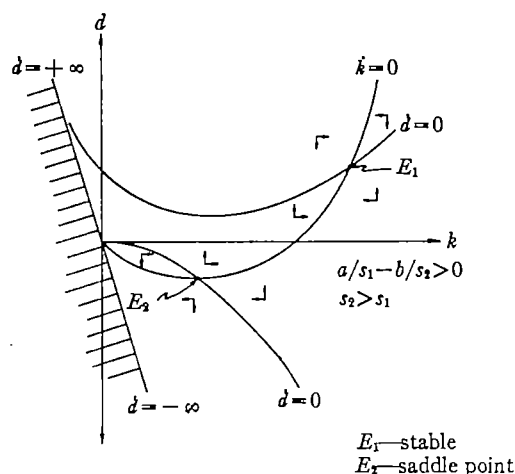


FIGURE 2

repays constant fractions of its national income. For the capital exporter, these fractions—representing annual foreign investment and service receipts from abroad—are constant as well.

REFERENCES

1. P. K. BARDHAN, "Equilibrium Growth in the International Economy," *Q.J.E.*, Aug., 1965, pp. 455-64.
2. K. HAMADA, "Economic Growth and Long-Term International Capital Movements," *Yale Economic Essays*, Spring, 1966, pp. 49-96.
3. H. JOHNSON, "Economic Expansion and International Trade," *The Manchester School of Economic and Social Studies*, May, 1955, pp. 95-112.
4. H. ONIKI AND H. UZAWA, "Patterns of Trade and Investment in a Dynamic Model of International Trade," *Rev. of Econ. Studies*, Jan., 1965, pp. 15-38.
5. J. H. WILLIAMS, "The Theory of International Trade Reconsidered," *Econ. J.*, June, 1929, pp. 195-209.

APPLICATIONS OF ENGINEERING ANALYSIS OF PRODUCTION TO ECONOMETRIC MODELS OF THE FIRM*

By ROBERT J. ANDERSON, JR.

Purdue University

I

Although Chenery [3, p. 531] has suggested that the purely statistical approach to determination of production, cost, and factor demand relationships, and the engineering approach (e.g., [3] [9] [11]) are complementary, subsequent practitioners of the engineering approach have used market data only to determine nontechnological parameters needed to calculate economic magnitudes from engineering relationships. No attempt has been made to test rigorously the results of engineering calculations against market data or to otherwise incorporate engineering information into statistical model building and testing procedures.

This paper suggests a synthesis of information derived from engineering theory with the accounting data from which econometric models of the firm are generally constructed. In instances where the behavior of the firm is strongly modulated by technology, engineering data can convey a considerable amount of prior information about costs and factor demands. Moreover, the use of such prior information can materially improve the statistical resolution of the model.

In Section II, an approach to synthesizing engineering theoretic information and sample data will be outlined. Special attention is devoted to the statistical desirability of the suggested synthesis. Section III presents an application of the suggested approach. In Section IV, some concluding observations are offered.

II

In most general terms, synthesis of engineering and accounting data is effected in three stages: (1) The heuristic constrained cost-minimization problem, which is the source of qualitative theorems on costs and factor demands, is replaced by a representation of the engineering constrained cost-minimization problem which confronts the firm. (2) The engineering cost-minimization problem is solved for various constraints on output(s) and various factor price combinations.

(3) The solutions thereby derived are used to restrict the maintained hypothesis under which estimation and testing, predicated on sample data, are conducted.

In practice, the synthesis is seldom achieved as easily as this general three-stage summary would suggest. A host of problems, some practical and some conceptual, arise in formulating and solving the engineering cost-minimization problem. Many of these specific difficulties have been discussed elsewhere [3] [9]. The more general difficulties which tend to be associated with practical mathematical programming problems have also received extensive consideration [6]. Attention here is therefore directed to the third stage of the synthesis.

Solutions of the engineering-cost model for alternative output constraints and factor prices yield data on long-run total costs, marginal costs, and factor demands. Because the engineering model and factor prices associated with engineering variables are likely to be uncertain in some degree, solutions are probably best treated as if they were random. The distribution of solutions, however, cannot in general be deduced from the distribution of technological and factor price parameters. Virtually nothing is known about the propagation of randomness through programming models [6, p. 18].¹

Nonetheless, engineering model solutions are broadly suggestive of the form of cost and factor demand relationships and can be fitted by numerical methods to obtain approximations to these relationships. If, in addition, one wishes to make some heroic assumptions about the distribution of model solutions, these approximating forms may be given a probabilistic interpretation.

The use of the information embodied in these approximations, be it given an exact or a probabilistic interpretation, can make several theoretically beneficial contributions to the estimation of model parameters. To the extent, for example, that such information aids in specifying the form of some equations and the variables to be included therein, it adds assurance that model parameter

* Helpful comments and suggestions from Lawrence Klein and Almarin Phillips bearing on many aspects of the dissertation from which this paper is extracted are gratefully acknowledged.

¹ Although some simple engineering models have been expressed as classical optimization problems (e.g., [11]), the more general programming problem formulation is assumed in all the discussion that follows.

estimates will be consistent and asymptotically unbiased [2]. Identifying information is also gained.

Perhaps most important, the incorporation of information derived from engineering theoretic considerations directly into the estimation procedure results in more efficient estimation of model parameters. Although it is generally accepted that the greater the amount of information brought to bear on an estimation problem, the greater the efficiency of estimation, extant proofs of this proposition (e.g., [8] [5, p. 257]) are established under unnecessarily restrictive conditions. A more general proof, encompassing linear and nonlinear models subject to linear and/or nonlinear restrictions, either probabilistic or exact, is outlined below for a normal single-equation model. Asymptotic results are derived here since small sample results are established only under fairly restrictive conditions.

Consider the hypothesis²

$$\begin{aligned} y_i &= f(x_i | \theta_1, \dots, \theta_r) + \epsilon_i \\ g_i(\theta_1, \dots, \theta_r) &= 0, \quad i = 1, \dots, m < r \\ \epsilon_i &\sim N(0, \sigma^2), \quad i = 1, 2, \dots, T \end{aligned}$$

where $f(\cdot)$ and the $g_i(\cdot)$ are presumed to possess bounded derivatives with respect to the vector θ up to the third order. It is also assumed that there exists at least one nonsingular $m \times m$ submatrix of the Jacobian, ∇G , of the $g_i(\cdot)$, $i = 1, \dots, m$, in the neighborhood of θ^0 , the true value of θ . Without loss of generality let the first m columns of ∇G be linearly independent in the neighborhood of θ^0 , and denote this submatrix by ∇G_m . Then, there exists a representation of $\theta_D = (\theta_1, \dots, \theta_m)$ in terms of $\theta_r = (\theta_{m+1}, \dots, \theta_r)$ in the neighborhood of θ^0 . Let L be the log of the likelihood function defined for the r -dimensional vector θ and let L^* be the log of the likelihood function defined for the $(r-m)$ -dimensional vector θ_r .

As a necessary condition that L obtain a maximum subject to the constraints,

$$\frac{\partial L^*}{\partial \theta_r} = \frac{\partial L}{\partial \theta_r} - \nabla G'_{(r-m)} \lambda = 0$$

$\nabla G_{(r-m)}$ is the submatrix of ∇G formed by deleting the first m columns and λ is the vector of Lagrangian multipliers.

Under the restrictions on $f(\cdot)$ and the $g_i(\cdot)$, $i = 1, \dots, m$, the asymptotic covariance matrix

² In this context, the normal single-equation model specification is not restrictive since under most general specifications of the random structure of a model, errors are asymptotically normal. Extension to multiple equation models requires only a proliferation of algebra.

of the maximum likelihood estimator of θ_r is given [4, p. 502] by

$$\begin{aligned} \text{var}^*(\theta_r) &= - \lim_{T \rightarrow \infty} E \left\{ \frac{\partial^2 L^*}{\partial \theta_r \partial \theta_r'} \right\}^{-1} \\ &= \lim_{T \rightarrow \infty} E \left\{ \left(\frac{\partial L^*}{\partial \theta_r} \right) \left(\frac{\partial L^*}{\partial \theta_r} \right)' \right\}^{-1} \end{aligned}$$

Then

$$\begin{aligned} \lim_{T \rightarrow \infty} E \left\{ \left(\frac{\partial L^*}{\partial \theta_r} \right) \left(\frac{\partial L^*}{\partial \theta_r} \right)' \right\} \\ &= \lim_{T \rightarrow \infty} E \left\{ \left(\frac{\partial L}{\partial \theta_r} \right) \left(\frac{\partial L}{\partial \theta_r} \right)' \right\} \\ &\quad - \lim_{T \rightarrow \infty} E \left\{ \left(\frac{\partial L}{\partial \theta_r} \right) \lambda' \nabla G_{(r-m)} \right\} \\ &\quad - \lim_{T \rightarrow \infty} E \left\{ \nabla G'_{(r-m)} \lambda \left(\frac{\partial L}{\partial \theta_r} \right)' \right\} \\ &\quad + \lim_{T \rightarrow \infty} E \left\{ \nabla G'_{(r-m)} \lambda \lambda' \nabla G_{(r-m)} \right\} \end{aligned}$$

Noting that

$$\lambda = (\nabla G_m^{-1})' \left(\frac{\partial L}{\partial \theta_D} \right),$$

and letting

$$M = \text{plim}_{T \rightarrow \infty} \{ \nabla G'_{(r-m)} (\nabla G_m^{-1})' \}$$

the following result is obtained.

$$\begin{aligned} [\text{var}^*(\theta_r)]^{-1} &= [\text{var}(\theta_r)]^{-1} + [I \mid -M] \\ (1) \quad &\begin{bmatrix} 0 & [\text{covar}(\theta_r, \theta_D)]^{-1} \\ [\text{covar}(\theta_D, \theta_r)]^{-1} & [\text{var}(\theta_D)]^{-1} \end{bmatrix} \begin{bmatrix} I \\ -M' \end{bmatrix} \end{aligned}$$

The inverse of the asymptotic covariance matrix of the constrained estimator may thus be expressed as the sum of a submatrix of the inverse of the asymptotic covariance matrix of the unconstrained estimator and a positive semidefinite matrix, the elements of which are transforms of the elements of a submatrix of the inverse of the asymptotic covariance matrix of the unconstrained estimator.

Consider first the effect of imposing restrictions on the efficiency of estimation of a single element of θ_r , say θ_{m+1} . Since the maximum likelihood estimator (restricted or unrestricted) is asymptotically normally distributed, it follows that the asymptotic variances and covariances of individual elements of either estimator may be directly obtained from the asymptotic covariance matrix for the joint distribution of the corresponding estimator. As a corollary, it also follows that the inverse of the asymptotic covariance matrix of marginal distributions for either esti-

mator may be obtained by inverting the relevant submatrix of the asymptotic covariance matrix of the joint distribution. Letting η be the inverse of the asymptotic variance of the restricted estimator of θ_{m+1} , letting $\gamma > 0$ be the corresponding inverse for the unrestricted estimator of θ_{m+1} , and $\delta \geq 0$ be the corresponding element of the second matrix on the right-hand side of (1) above,

$$\eta = \gamma + \delta, \quad \eta \geq \gamma, \quad \text{which implies} \quad \frac{1}{\eta} \leq \frac{1}{\gamma},$$

that is, the constrained estimator of θ_{m+1} has smaller asymptotic variance than the unconstrained estimator. Since numbering of the elements of θ_I is arbitrary, this result clearly holds for all elements of θ_I .

Comparing generalized asymptotic variances of the constrained and unconstrained estimators

$$\det \left[\text{var}_{T \rightarrow \infty} (\theta_I^*) \right]^{-1} \geq \det \left[\text{var}_{T \rightarrow \infty} (\theta_I) \right]^{-1}$$

since

$$\left[\text{var}_{T \rightarrow \infty} (\theta_I^*) \right]^{-1} \text{ is the sum of } \left[\text{var}_{T \rightarrow \infty} (\theta_I) \right]^{-1}$$

and a positive semidefinite matrix, which implies that

$$\det \left[\text{var}_{T \rightarrow \infty} (\theta_I^*) \right] \leq \det \left[\text{var}_{T \rightarrow \infty} (\theta_I) \right]$$

as was to be established.

If the constraints $g_i(\cdot)$ $i=1, \dots, m$ are treated as asymptotically normal variates distributed independently of sample data and the estimation problem is treated as one of maximizing the joint likelihood function of the sample and the constraints,³ an exactly analogous result is obtained except for the fact that the imposition of probabilistic restrictions does not reduce the dimension of the parameter space.

III

An econometric model of bituminous concrete firms is the vehicle for a brief demonstration of the practical properties and problems associated with the proposed synthesis.⁴ Only the cost side of the model will be considered here.

Bituminous concrete (variously termed "asphaltic concrete," "hot mix," and "hot asphalt") is the familiar blacktop paving mixture found on many highways, streets, parking lots, and so on. In barest terms, its production consists of heating

and drying mineral material, dubbed "aggregate" and mixing the hot dry aggregate in sized proportions by weight with hot asphalt (the petroleum distillate). The production of hot mix involves the concomitant production of an unwanted by-product, great quantities of particulate air pollutants (dust). The firm was thus treated as a joint product enterprise, producing both bituminous concrete and, viewing the physical output of particulates negatively because of utility considerations, reductions in emissions of particulate pollutants.

The typical plant may be divided into three processes: (1) cold feed, (2) drying (and exhaust), and (3) mixing. Efficient production requires strict technical balance among the three processes. Of the three, only the drying and exhaust operation has interesting possibilities for substitution among engineering inputs. In addition, the drying process encompasses the principal air pollution control operations of the plant and is the most important determinant of long-run costs. Therefore, only the drying process was analyzed in detail.

The engineering literature on the drying process, including air pollution control equipment, was distilled into a nonlinear nonconvex programming problem in twelve variables—dryer diameter, dryer length, volume rate of flow of dryer gases, dust collector diameter, dust collector height, fuel use, number of dust collectors, aggregate use, dryer slope, pressure drop across a scrubber dust collector, pressure loss at atomization in scrubber liquor, and scrubber liquor flow rate—and twelve nontrivial constraints.

The model was solved for a number of combinations of constraints on emissions and hourly output capacity. Solution values for inputs closely approximated specifications for actual plants with output and emission reduction capabilities resembling those required by the constraints, and the values of the objective function at solution suggested average costs of drying near those cited as usual by industry representatives, approximately \$2.30 per ton. Although the programming model was nonconvex, choice of starting values for the nonlinear programming algorithm, which were, to be frank, near what I hoped would be the solution values, seemed sufficient to ensure convergence to a reasonable local minimum.⁵

³ This treatment may be interpreted as the solution to a stochastic programming (e.g., [6, p. 170]) problem wherein the objective function is the sample likelihood and the expected loss of violating the constraints is the likelihood of the constraints. The Bayesian regression estimator [10, pp. 334–53] is a special case.

⁴ The model and data sources are discussed in detail in [1].

⁵ The fact that only local minima may be obtained is of no consequence. The purpose of solving the engineering model is to explain what firms do, not what they ought to do. If firms find only local optima, these are the optima we wish to find.

Scatters of costs against output of hot dry aggregate and emission reductions suggested that (a) output and costs, with emissions held constant, are linearly related, with the intercept increasing slightly as emissions are reduced, and that (b) costs and emissions tend to be non-linearly related, perhaps as a second-degree polynomial. Such a function was fitted to the solution data by ordinary least squares, with the result

$$(2) \quad \frac{C}{1000} = \frac{1.1484}{(.0010)} \left(\frac{X}{494.25} \right) - \frac{.0188}{(.0044)} \left(\frac{X}{494.25} \right) \left(\frac{E}{1000} \right) + \frac{.0164}{(.0053)} \left(\frac{E}{100} \right)^2$$

$$R^2 \text{ (adjusted for degrees of freedom)} = .9999$$

where C is total annual cost of drying, X is total annual output in tons, and E is emissions in pounds per hour.

Under the generally reasonable assumption that costs of cold feed and mixing are proportional to output and additive to drying costs, a function of the form of equation (2) above was fitted by ordinary least squares⁶ to cross-section data for a sample of 27 bituminous concrete firms, and the following result was obtained

$$(2a) \quad \frac{TC}{1000} = \frac{5.5858}{(.1537)} \left(\frac{X}{1000} \right) + \frac{.2151}{(.3817)} \left(\frac{X}{494.25} \right) \left(\frac{E}{1000} \right) - \frac{.0513}{(.1031)} \left(\frac{E}{100} \right)^2$$

$$R^2 \text{ (adjusted for degrees of freedom)} = .9775$$

The coefficient of the total output variable in (2a) agrees closely with the pronouncements of industry experts on average long-run production costs. Less satisfactory are the coefficients associated with emissions. Although they confirm, in conjunction with the coefficients of equation (2), the notion that air pollution control costs are not substantial in relation to total costs, the statistical resolution of equation (2a) is unsatisfactory. This is not unexpected since it is generally difficult to obtain good estimates of cost functions for joint product firms, particularly where the impact of one output on cost dwarfs the other.

Treating the sample distribution of the co-

efficients of the pollution variables in (2) as prior probabilistic information on these model parameters, and weighting this information together with that of the sample in accordance with procedures established in [10], the following equation results.

$$(2b) \quad \frac{TC}{1000} = \frac{5.6435}{(.01008)} \left(\frac{X}{1000} \right) - \frac{.0188}{(.0044)} \left(\frac{X}{494.25} \right) \left(\frac{E}{1000} \right) + \frac{.0164}{(.0053)} \left(\frac{E}{100} \right)^2$$

$$R^2 \text{ (adjusted for degree of freedom)} = .9771$$

The gain in efficiency is immediately evident. Prior information virtually swamps the sample in determination of the posterior means of the coefficients of the emission variables. Tests also indicate that the hypothesis that sample and prior means are equal cannot be rejected at reasonable levels of significance; the resolution of the sample mean is simply too poor.

IV

Although improvement of the statistical properties of models has been stressed as the principal justification for incorporating engineering information into models of the firm, it is quite likely that the usefulness of such information can transcend rigorous statistical interpretation. One obvious merit of the method is its great power to ferret the costs associated with joint products even where the costs associated with one of the outputs are, relatively speaking, quite small.

Closely related to the problem of allocating costs between joint products are problems concerning the nature and role of technically non-productive activities of the firm. To the extent that engineering costs reflect adequately the costs associated with the physical transformation of inputs into outputs, differences between accounting costs and engineering costs may reflect such things as advertising outlays and extraordinary payments to some factors. It is generally not possible to unearth these nonproductive outlays from accounting data alone. Evidence thereon would materially aid efforts to investigate the behavior of firms, and to relate that behavior to market structure or other organizational characteristics.

A fundamental advantage conveyed by the use of engineering information is that its autonomy [7, p. 84] can be exploited to yield information about structural changes long before such changes may be adequately estimated from accounting

⁶ Output and emissions are assumed to be exogenously determined in the cross-section, determined principally by governmental expenditures in the locale served by a firm and governmentally imposed pollution regulations.

data, and perhaps long before such changes begin to be manifested therein.

Use of engineering analysis can also free the investigator from the constraints of limited empirical experience in another way. It frequently happens that in sample periods of short duration or limited enough scope so that observations may reasonably be assumed to have come from the same random structure, the range of observed variation is not very broad. This tends to give undue weight to observations which represent only a part of the structure in determining the whole structure. For example, Klein argues that Dean's finding of linearity in total cost functions may well be a result of excess capacity during the sample period, and that nonlinearities might appear at higher rates of utilization [7, p. 121]. If an engineering model is available, however, constraints and parameters can be varied at will. Engineering information thus gives one a kind of clairvoyance, an ability to see beyond the confines of what is frequently a very limited sample.

REFERENCES

1. R. J. ANDERSON, JR., "Application of Engineering Analysis of Production in Econometric Models of the Firm" (unpublished Ph.D. dissertation, Univ. of Pennsylvania, 1969).
2. T. W. ANDERSON AND H. RUBIN, "The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Equations," *Am. Math. Stat.*, Dec., 1950, pp. 570-82.
3. H. CHENERY, "Engineering Production Functions," *Q.J.E.*, Nov., 1949, pp. 507-31.
4. H. CRAMER, *Mathematical Methods of Statistics* (Princeton, 1946).
5. A. S. GOLDBERGER, *Econometric Theory* (Wiley, 1964).
6. G. HADLEY, *Nonlinear and Dynamic Programming* (Addison-Wesley, 1964).
7. L. R. KLEIN, *An Introduction to Econometrics* (Prentice-Hall, 1962).
8. ———, "The Efficiency of Estimation in Econometric Models" in R. W. Pfouts, ed., *Essays in Economics and Econometrics in Honor of Harold Hotelling* (Univ. of North Carolina, 1960), pp. 216-32.
9. A. S. MANNE AND H. M. MARKOWITZ, eds., *Studies in Process Analysis* (Wiley, 1963).
10. H. RAIFFA AND R. SCHLAIFER, *Applied Statistical Decision Theory* (Harvard, 1961).
11. V. L. SMITH, *Investment and Production* (Harvard, 1961).

DISCUSSION

LIONEL W. MCKENZIE: It is an interesting project to introduce stochastic elements into the equations of a Walrasian tatonnement and then to examine conditions that will guarantee some form of stability. Also, it is worthwhile to compare conditions that guarantee stability when the stochastic elements are present with conditions that guarantee stability when they are absent. Thus I will be happy to see Mr. Weintraub and his former colleague Mr. Turnovsky continue their work in this direction. In order to give them a modest assist in this project I will make a few suggestions, most of which will already have occurred to them.

First, however, I would ask for caution in interpreting the results that may come. For example, if we find that the models of the competitive adjustment process that we construct are stable only under difficult assumptions, this does not imply that markets are not competitive. It only implies that if the markets are competitive, our models of the adjustment process are poorly chosen. On the other hand, if we find that the conditions we can discover to provide stability for the stochastic model are stronger than the conditions we have found that stabilize the deterministic model, this does not imply that the stochastic model is less stable. Indeed, it does not even imply that the conditions that stabilize the deterministic model do not stabilize the stochastic one as well. To prove that we must provide appropriate examples of stochastic models which are unstable and which satisfy the stability conditions for the deterministic case. Actually this should usually be easy to do.

More constructively, I would urge as a general methodological rule that one try very hard to introduce into the stability conditions as much as possible that is peculiar to the economic problem. For example, in most of the competitive stability models prices are confined to the positive orthant. Thus it is natural from the economic viewpoint to use the positive orthant as the neighborhood of equilibrium within which the stochastic price path can wander. It is odd to choose a symmetric neighborhood of the equilibrium point which is the largest that can be contained in the positive orthant and use an assumption on the stochastic process just for that region. This results in a theorem in which this odd neighborhood plays a critical role. What one should do is to return to the proof of the stochastic theorem on which you depend and see if there is anything in the mathematics which prevents the use of a more natural region from an economic viewpoint. One cannot expect the mathematician to have anticipated the special features of the economic problem. These features may justify one in designing a stochastic process with barriers at the boundaries of the positive orthant. I would wonder if the probability of convergence to equilibrium in Theorem 3 could not be raised to one by some such revisions.

Another move which it would seem natural to make is to allow the stability condition, that is, the condition for having a Liapounov function which is a supermartingale to fail in a smaller neighborhood of equilib-

rium, where excess demand is small but errors persist, and then prove a theorem that sends the path to this neighborhood of equilibrium with some probability, perhaps equal to one.

Now I would like to list some specific questions to the author on points in his paper which I failed to understand completely. He says relative prices are given initially. I presume he means some actual absolute prices. Then in the deterministic process the length of the price vector is constant and convergence occurs to a particular equilibrium price vector. However, in the stochastic process I suppose this is no longer true so convergence must be to a point on the equilibrium price ray. This homogeneity of the equilibrium price set is peculiar to the economic case, the more reason to give it careful treatment.

The author gives a condition on the infinitesimal operator in Theorem 3 that cannot hold unless the power of the disturbance approaches zero near equilibrium. One would like to see a discussion of the plausibility of the disturbance taking this form.

The author may have been a little too brief in his discussion of existence of equilibrium. In order for continuous demand functions to be consistent in all cases with utility maximization some assumptions are needed. It will do to have insatiability in all goods individually and all consumers holding positive stocks of all goods.

LOWELL S. HARDIN AND G. E. SCHUH: We congratulate Professor Belli for going about his research in an imaginative manner. Empirical work on the development process is important. Especially needed is rigorous testing of relevant hypotheses concerning human behavior.

Professor Belli states that on the surface it appears that Nicaraguan farmers have behaved irrationally. They increased their cotton output despite what appeared to be negative price and income incentives.

But, says Professor Belli, this really is not what happened. Cotton farmers had constant returns to scale; unit costs may have declined a bit due to technological changes; the farmers learned by doing, but there was a time lag; and credit was liberalized permitting freedom of entry if one could demonstrate reasonable skills in cotton production.

Reading Professor Belli's paper, one has the feeling that he gets only part of the story. Perhaps the missing insights are discussed in the larger study on which this paper was based.

Let us speculate. Could the Nicaraguan cotton case have unfolded somewhat as follows?

Early on in the shift to greater cotton production, prices were high and margins wide. Nicaragua developed firm market outlets for its high quality cotton, primarily with Japan. While prices fluctuated around world levels, a solid marketing system involving grades, standards, storage and assurance of payment was developed. Further, cotton as compared to other enterprises received indirect and direct subsidies. Adequate credit at what were probably subsidized terms was pro-

vided. The banks put on field men to spread the technology and serve as trouble shooters, since no overall extension system existed. Price, market and cash flow as well as technical uncertainties were reduced.

Given this package, farmers shifted out of other crops, increased their cotton acreage, and became more specialized producers. Even though cotton prices fell, superior alternatives did not appear. Investments in learning and in specialized equipment had been made. As viewed by the farmer, he was better off to raise cotton even at the lower prices than to divert his resources to other uses. Yet cost-price relationships, at least insofar as variable costs were concerned, were such that he was better off to farm than to go out of business. Nonfarm alternatives, given his capabilities, were not attractive either.

Moving from speculation to specifics, Professor Belli's analysis of technical change would be made more convincing were he to provide data on varieties, fertilization, plant protection, and yield per hectare. Was there a net flow of resources from the nonfarm sector? How severe was capital rationing? What were the real interest rates? Was capital equally available for other enterprises?

Summarizing, Professor Belli has given us a useful paper. He has excited our curiosity as evidenced by the above questions. His introduction of learning theory into the analysis is a thoughtful step away from usual static models.

I for one, however, am willing to put to rest the issue of price and income responsiveness of farmers. There are more important development problems begging the attention of competent researchers. Take the Nicaraguan cotton case as an example. A reported 250 mechanical cotton pickers were imported in the thrust for accelerated production. Given the harvesting losses and the relative prices of capital and labor, producers are said to be reverting to hand harvesting. One wonders whether at some stage the cost of capital (as represented by the mechanical pickers) was priced at unrealistically low levels relative to the price of labor. Differential exchange rates may have been used. Following such signals, the farmer maximized the returns to his firm by specializing and mechanizing. But this same decision-making process may have materially lowered the social return on the sizable investments involved.

Given the pressure for employment generation which characterizes developing economies, one wonders whether sufficient attention is given to labor versus capital-intensive technology and policies? This, of course, is but part of the needed larger analysis of the whole labor market.

To what extent is uncertainty rather than absolute cost-price relationship a dominant factor inhibiting productivity advances? More usefully, what feasible technological and policy changes can man devise to cope with uncertainty issues?

Finally, is there a chance that developed and developing nations can, in this decade of nationalism, re-examine trade and marketing policies and practices?

Again in the Nicaraguan case we speculate that the marketing tie to Japan had much to do with what has transpired. If elbow room exists for more research on agricultural products which enter international trade, where do we start? Are we boxed in by the self-interests of dominant nations, or is there a development pay-off in shorter-run trade concessions? If significant policy changes are to be made, the developed nations will have to take the lead. A necessary prerequisite is empirical analysis of the probable consequences.

RONALD W. JONES: Professor Myers' paper provides a neat and compact model of a two-country trading world with international capital mobility. One of the simplifying features of the model is the complete separation of productive activity between the two countries: Agraria (country 1) is the sole producer of the world's supply of consumer goods while Mancunia (country 2) is specialized as the only producer of infinitely long-lived capital. No question of the comparability of technologies need arise since no goods are produced in common. "Capital movements" ensure that rates of return to capital are equalized between countries, and the mathematical properties of the model (greatly simplified by the assumption of Cobb-Douglas technology) are exploited by Myers to develop two differential equations for the time derivatives of the world capital stock and the "debt" position. The use of phase diagrams allows Myers to explore the convergence to balanced growth properties of the model.

In commenting upon this paper I do not intend to challenge the results Myers obtains. According to my calculations he has correctly dealt with the model. Instead, I wish to remark, first, on the nature of the capital movements involved and, second, highlight one of the crucial comparisons between countries that involves savings ratios and capital coefficients.

The only way that Mancunia can obtain consumption goods is by exporting capital goods. In this sense the existence of trade in this model is synonymous with a type of capital movement. Nonetheless, it could be required that these current account flows balance, and later I shall briefly explore the model at that stage. However, Myers allows, in addition, for a debt position to be built up. One could imagine one of the countries anxious to "live beyond its means" and allowed to do so by a willingness on the part of the other country to refrain from currently consuming and investing an amount equal to its national income. The extent of international borrowing and lending would then be conditioned by behavior on the demand side, which might be sensitive to prevailing rates of interest. The point I want to make about Myers' model is that this is not the type of capital movement he is allowing. Rather, the capital flow is a response to differentials in rates of return to capital in the two countries and is of the type familiar to the literature on foreign investment. The receipts from Agraria's current exports of consumer goods are used by Agraria's residents to purchase capital goods from Mancunia. An additional capital flow would take place if, say, the return to capital in Agraria is higher than in Mancunia. But this would not represent

sales of capital to Agraria. It would be reflected in the shipment of capital goods to Agraria by residents of Mancunia, who repatriate the income earned by that capital, part of which is consumed $(1-s_2)$ and part invested.

The other point I wish to make is related to this. A crucial term in Myers' differential equation (14) for the time change of the "debt" position is his constant C_3 , the sign of which depends upon a comparison of s_1/a with s_2/b , where a and b are the capital shares in 1 (Agraria) and 2 (Mancunia), respectively. By considering Myers' model in the absence of any foreign investment it is possible to relate the comparison of s_1/a with s_2/b to the pre-foreign investment spread in rates of return to capital. If no foreign investment is allowed, the value of Agraria's exports, s_1PX (the nonconsumed part of its output), would equal the value of its imports, $(1-s_2)Z$ (that part of Mancunia's capital production not used for local investment). That is,

$$P = \frac{(1-s_2)}{s_1} \frac{Z}{X}$$

Let r_1 and r_2 denote each country's rate of return on capital (the rental deflated by the price of the capital good). Then

$$r_1 = a \frac{PX}{K_1} = \frac{a(1-s_2)}{s_1} \frac{Z}{K_1}$$

$$r_2 = b \frac{Z}{K_2}$$

And Agraria's rate of return to capital would be higher than Mancunia's if and only if

$$\frac{a(1-s_2)}{s_1} > b \frac{K_1}{K_2}$$

or:

$$\frac{K_1}{K_2} \frac{s_1}{a} < \frac{1}{b} - \frac{s_2}{b}$$

Adding s_1/a to both sides and rearranging:

$$\left(\frac{s_1}{a} - \frac{s_2}{b}\right) > \frac{K}{K_2} \left(\frac{s_1}{a} - \frac{K_2}{bK}\right)$$

where K is the world capital stock, K_1+K_2 . But if no foreign investment is allowed, in a steady state the ratio K_2/K must reflect Mancunia's savings ratio, s_2 , since K represents past accumulated outputs of Mancunia (which only produces capital). Therefore in the pre-foreign investment stage,

$$r_1 \geq r_2 \quad \text{if and only if} \quad \left(\frac{1}{s_2} - 1\right) \left(\frac{s_1}{a} - \frac{s_2}{b}\right) < 0$$

Thus the direction of foreign investment, once it is allowed, is towards Agraria if and only if the savings coefficient, deflated by the capital coefficient, is larger in Mancunia. Incidentally, this argument does not depend upon a and b being the capital exponents in Cobb-Douglas functions—they need only be capital shares. Finally, note that in Figure 2 in Myers' paper this suggests that if $s_1/a < s_2/b$, as he there assumes, a long-run equilibrium solution with foreign investment must involve $D > 0$ (i.e., Mancunians investing in Agraria); that is, this supports the stable equilibrium E_1 and rules out the saddle-point possibility, E_2 .

INCREASING THE SUPPLY OF BLACK ECONOMISTS IS ECONOMICS CULTURE-BOUND?

By KENNETH E. BOULDING
University of Colorado

In one sense, economics, like any discipline, is culture-bound almost by definition, because it is the product of the subculture of economists—a subculture which is differentiated from other subcultures and from which economists could not escape without ceasing to produce economics. If the question, “Is economics culture-bound?” is to have any meaning, it must be interpreted to signify, “Does the character of economists and the nature of their subculture affect the scope, the relevance, or the truth of the product in the shape of economics which they produce?” Stated in this way, the question is clearly important, and if answering it leads to increased self-consciousness among economists about the nature of their own subculture and their own product, this may change the direction and the content of economics itself.

The subculture of economists, as it has existed now for nearly two hundred years, is very far from being a random sample of human populations or human cultures. It bears the genetic stamp of its British origin, of its eighteenth-century birth date, of its connection with the larger scientific subculture, and of its origins in what might be called the intellectual middle class. The important question is would economics have been any different if it had originated, let us say, in China or in India, or if its practitioners were less middle class, or if the economic structure of the profession, divided as it is between education, business, and government institutions, have been different? These are questions perhaps to which only speculative answers can be given. Nevertheless, the questions themselves are worth raising, if only because they may point to areas where there may have been a misallocation of intellectual resources in economics, or where economists have taken things for granted that they should not have taken for granted, or have limited their scope and interests because of the scope and interests of their own subculture. The application which is of particular interest to this symposium is, of course, whether the class, race, or national affiliation of economists has led to a neglect or distortion of the subject matter and to the neglect of problems which are of particular interest to disadvantaged groups, defined by skin color or by cultural, lin-

guistic, religious, or other factors. On the other side, does economics tend to become an apologia for a *status quo* which perpetuates poverty, discrimination, and the class structure of society? The answer to these questions may well lie in an examination of the impact of economics on economics and social policy. What, for instance, have been the results for society, and especially for the poor and disadvantaged, of those social changes which can be attributed to economics? It is by our successes that we may have to be judged!

Any answers which are suggested in this paper must be regarded as highly tentative, for the major function of the paper is to raise questions rather than to give answers. What follows expresses a tentative point of view consisting more of hypotheses than of well-established propositions.

The first proposition is that economics is perceived to be defective if it is in fact culture-bound in the way outlined above. Economics is a subset of human knowledge about the sociosphere; that is, about the total social system of the world. It deals particularly with those aspects of the sociosphere and human relationships and organizations which are organized through exchange and through the transfer of exchangeables. I prefer this definition, incidentally, to a common definition that economics is the study of the allocation of scarce resources, simply because there are other methods of allocation of scarce resources which do not involve exchange or exchangeables, and which are more in the province of political science or sociology. Wherever there is exchange, therefore, there is something to know about it, and what there is to know about it is economics. The concept of the exchangeable is perhaps even more fundamental here than that of exchange, for economics deals with the production and the consumption of exchangeables, as well as with exchange, and it also deals with the one-way transfer of exchangeables in the grants economy. Exchange is found very early in the development of human society. There may indeed be some extremely primitive societies in which it barely exists and in which all allocations are made by threat or by traditional authority. The history of “silent trade” at least suggests that exchange is not always easy to legitimate. Exchange seems to

take an increasingly important role as society develops, though we do find periodic revolts against it and attempts to create societies which are based more on acceptance of status, persuasion, integrative relationships, or on threat rather than on exchange.

Poverty is a matter of provisions rather than of exchangeables, as such; that is, it is a matter of the "necessities, conveniences, and luxuries of life," which a person or a family can command and enjoy, as Adam Smith taught so clearly. It could be argued indeed that economics began because of an interest in poverty and wealth, that is, in provisions, and that it was only when it became clear that most provisions (goodies) were, in fact, exchangeables (goods) and were provided through exchange that economists got interested in exchange itself. It could be argued indeed that this interest in exchange eventually overweighed the interest in poverty and wealth, and that economists lost interest in poverty as they gained interest in exchange. One sees this even in Ricardo, who is fascinated by the problem of the relative price structure, which is, of course, a problem of exchange, and which is what he means by the problem of "value." He only incidentally notices that value and riches are not the same thing at all. He is certainly less clear than Adam Smith on how the exchange system, in fact, produces riches. In the neoclassical economics of Marshall, also, and indeed in virtually all his contemporaries, even the mathematical economists, the primary focus of interest is the exchange system or the price system which so largely governs it, and not on development, which is the movement from poverty into riches. It is only in the last generation that economists have returned to the problem of poverty and riches and regained even the interest in this subject which Adam Smith had. It could be argued, therefore, that it was not so much the class, nationality, or race of economists that produced this effect, but the sheer practice of economics itself. The questions which are interesting to scientists are not always those which are most relevant to society and economics perhaps fell into the intellectual trap of being fascinated with the intellectually fascinating, perhaps at the cost of losing interest in things which were more important.

It has been argued, also, with some cogency that the time, place, and class of the economists did affect their concepts and their interests and the kind of models of society which they constructed. It has frequently been argued, for instance, that the classical division of factors of production into land, capital, and labor reflected

the peculiar social circumstances of eighteenth- or nineteenth-century Britain with its land-holding aristocracy, its rising merchant-capitalist middle class and its rather clearly defined lower class of workers. These three classes were separated quite sharply from each other by forms of speech, manners, religion, and other cultural characteristics, and hence it is not surprising that the classical economists perceived them as associated with three different and independent factors of production. If economics had grown up in India, there might have been a factor of production for each caste. On the other hand, I think one can defend even the classical economics against the charge of being merely culture-bound. It could be argued indeed that the culture-boundness in this case was a positive advantage in that it enabled the British economists to perceive crude but fundamental divisions in society which were somewhat masked in other societies. In a similar way it might be argued that a society whose experience with falling bodies was confined to leaves, feathers, and snowflakes would never have discovered the law of gravity, and that it took a society where one could drop cannon balls from high towers, even shall we say from leaning towers, in order to separate the brute facts of gravitation from the delicate and intricate impact of air resistance. I am not certain indeed that the law of economic rent is as secure, as universal, or as significant as the law of gravity, but it does have somewhat the same characteristics, in that it is a crude structural phenomenon which underlies an enormous variety of cultural and economic forms.

Another possible test of culture-boundness of economics is to look at the impact which it has had on human behavior and especially on political policy. Until perhaps very recently economics has not had much effect on human behavior in general. It has only been with the development of linear programming and operations research that ideas which have come essentially out of economics have had very much effect on the behavior of businesses. Economics has had an effect, however, almost since the eighteenth century on government policy and especially on what was regarded as the ideals of government policy. Thus, the hesitant movement toward free trade since the eighteenth century, such as the abolition of the Corn Laws in England, and even a good part of the commercial policies of the United States today owe a good deal to the analyses of economists as far back as Adam Smith's great polemic against mercantilism. The case for free trade rests more on static analysis than it does on dynamic and it has been argued that whereas in the case of the

rich countries, from which the economists mostly came, free trade would have a continued favorable dynamic effect, in the case of the poor countries, which were ambitious for development, free trade may easily produce too much specialization which might pay off in the short run but would be a long-run impediment to the overall growth of the economy. The development of plantation-type economies in the tropics is a good case in point. In this case, therefore, it could well be that the fact that economists on the whole have come from rich and dominant countries may have diverted their attention from the problems of the development of poor countries and, hence, may have induced them to give advice which may have been better for the rich than for the poor.

Another achievement of the economists was the English Poor Law of 1834, inspired largely by Malthusian principles, whereby the poor were segregated into work houses and life was made pretty uncomfortable for them. This may well have been an example of failure of empathy on the part of economists. Classical economists, especially, were well meaning and rather gentle people who, unlike the mercantilists, certainly had the interests of the poor at heart. They had indeed a terrifying vision of the impossibility of getting rid of poverty in the absence of extravagant tastes and family limitation on the part of the poor, but their message was that, if anything, "if you want to get rid of poverty this is the only way to do it." It could well be that if the economists had had a little more personal experience in the culture of poverty that they would have realized that making paupers miserable was a very poor way of controlling the population, and it is highly probable that if the Poor Law had been less rigorous in England, the poor would have been a little happier, and nobody else would have been much the worse.

We can perhaps detect a class bias here. Adam Smith, for instance, speaks in one sentence of "the laboring poor and the laboring cattle," and educated people, well into the nineteenth century, did believe that the working class was almost another species who could stand misery much better than the refined classes. This attitude indeed persisted in Europe right into the first World War in the difference in attitude towards the enlisted men, whose sufferings far exceeded those imposed by the Poor Law, and the officers for whom life was frequently modestly comfortable. The idea that working people are people is a relatively modern one, and while it is hard to assess the contribution of economists to the spread of this idea, this contribution is by no means negligible. Cer-

tainly Adam Smith's emphasis on per capita real income as the major measure of national wealth leads directly to the notion that all "capitas" have some right to income and that the working class is not to be used simply as a means for producing national power or economic surplus. Still it was the Christians and not the economists who passed the Factory Acts and evangelical warmth probably had more to do with the development of a more integrative society than the chill and rather dismal models of the economists.

Another political device for which economists have a good deal of responsibility is the antitrust laws of the United States. These arise out of a rather abstract vision of an ideal system of perfect competition which the laws themselves quite fail to realize. Nevertheless, I am inclined to give them a credit balance. Whether they really directly help the poor very much may be doubted, but it is quite possible that the fear of the antitrust acts, and also such corresponding things as public utility regulation, induced in American business a willingness, which their European counterparts did not quite have, to invest and innovate rather than exploit market power. If, then, the antitrust laws contributed to development, we must certainly put them on the positive side, for development is by far and away the most efficient method for getting rid of poverty, although not necessarily for solving the problem of discrimination.

In the twentieth century the main impact of economics on economic policy has been the Keynesian Revolution and the development of full employment policies. The recognition that deflation is much worse than inflation and should be avoided at almost any cost came rather late, but it made a very important contribution to human welfare. Unemployment is such a major cause of poverty that the economists' contribution towards solving the unemployment problem is a large one that will pile up a large credit balance in any final judgment.

Economists also can take a good deal of credit for the movement towards more equal distribution of income and especially for the movement for doing this through direct payments, such as the negative income tax. Even though these movements have not produced very much in the way of legislation yet, there is a good chance they may do so in the future. The economists' perception that attempts at redistribution of income through manipulation of the price system are generally more likely to redistribute income to the rich than to the poor, as in the case, for instance, of our agricultural policy, is an important step in

undermining the more fraudulent aspects of social democratic legislation and in pointing out that the cheapest and most effective method of helping the poor is by direct payments.

In the last few years economists have shown a great deal of interest in the theory of discrimination, especially racial discrimination. Two principal figures in this field are Gary S. Becker, whose book, *The Economics of Discrimination* (University of Chicago Press, 1957), was perhaps the first major treatment of this problem from a theoretical point of view, and Lester C. Thurow, whose excellent study, *Poverty and Discrimination* (Brookings Institution), challenged Becker's models quite sharply. There is considerable divergence between these two authors, Becker regarding discrimination as a "taste" which has unfavorable consequences for both whites and blacks, whereas Thurow lays more stress on the monopoly power aspect of discrimination by which whites are able to have larger incomes than they otherwise would have. This is a problem of enormous complexity and it cannot be solved, of course, within the confines of economics, as it involves a very complex process of human learning and relationships. Economics, however, has something to contribute and has been particularly valuable perhaps in pointing up the real dimensions of the problem.

The principal failure of economics, certainly in the last generation, has been in the field of economic development, in spite of a great deal of attention which has been given to it. The refinements in development theory which have developed in the last generation do not seem to have carried us very far towards a real understanding of the process as a total social process, and we do not really understand what it is that makes the difference between a developing society and a nondeveloping society. Even within a developing society, like the United States, we do not really understand why some relatively poor segments of it, such as the blacks, on the whole are participating in the developmental process, though with about a thirty-year lag, while others, such as the Spanish-speaking Americans and Indians, hardly seem to be participating in the developmental process at all. The problems here seem to be beyond the power of economic models to reach and the most significant contributions come from social psychologists, such as McClelland, or psychological economists, such as Everett Hagen.

One wonders whether culture-boundness may not have something to do with this relative failure. Development, like economics, has been very largely a Temperate Zone product. The complexi-

ties both of tropical ecology and of tropical societies are beyond easy access for those raised in essentially Temperate Zone culture. This is not to suggest a naive climatological determinism, but just as tropical biological ecosystems differ very markedly from those in the Temperate Zone it would not be unreasonable to suppose that the processes of social evolution would likewise produce marked adaptations to the peculiar rigors and delights of tropical climate and life style. Culture-boundness is manifested primarily in the things that we take for granted. Thus, in the Temperate Zone we take for granted the rhythm of the seasons, with the constant looking forward into the future which this produces—"If winter comes, can spring be far behind?" The more disagreeable temperate climates produce demands for activity almost for its own sake. What else is there to do in the middle of the Chicago winter except work? The subtler rhythms and the more genial environment of the tropics may lead to a life style and a type of provisioning which is very different from that of the Temperate Zone.

The situation is complicated, of course, by the fact that, in the past at least, disease has been a major obstacle to development in the tropics, especially debilitating disease, for the environment which is agreeable to man is also, unfortunately, agreeable to his diseases. This source of differentiation has been diminishing with great rapidity with the conquest of tropical diseases and great irreversible changes may easily follow. Nevertheless, the economics of traditional tropical societies is still very different from societies of the Temperate Zone, whether socialist or capitalist, and it may well be that we need schools of tropical economics, just as we need tropical medicine, tropical ecology, and tropical agriculture. The result of imposing Temperate Zone techniques on the tropics, whether in engineering, agriculture, or in economics, may easily be disastrous. This perhaps is the point at which the culture-boundness of economics may be most significant.

We cannot leave this subject without recognizing that economics has two cultures at least, although which is the most culture-bound is hard to say. Marxian economics and radical economics generally form a separate stream which has common origins with capitalist economics in the classical economists, but which has diverged very strikingly both from its source and from its "bourgeois" relation. Karl Marx, of course, came out of much the same social class and background as the classical economists. Like Ricardo, his father was a Christianized Jew and the Judeo-Christian influence is very striking in the whole Marx-

ist system of thought. We can argue indeed that Marxist economics is even more culture-bound than capitalist economics, in that it carries with it the values and the culture of the European middle class. Indeed, the great split in the Communist camp between the Russians and the Chinese may reflect precisely this culture-boundness of Marxist economics.

The very existence of the Marxist split in economics, however, points up a still unresolved problem within the general body of the science, which arises perhaps because of the bias of classical and bourgeois economics towards individualism and towards the neglect of the economics of the group. We see this most clearly in things like utility theory and demand theory, in which a limited concession may be made to the family as a spending unit, but which basically regards the indifference curve, for instance, as a mental property of some individual. The factors of production, likewise, are individuals, not groups. The individual laborer decides how much labor to offer. The individual capitalist makes his individual decisions about how to invest his capital. The individual landowner decides what to do with his land. Group decisions and community demands and supplies are somewhat alien to the traditional framework of economics. Hence, the problems which revolve around group life, community, group identity, public goods, the grants economy, identity and identification, benevolence, malevolence, and so on have been neglected.

Part of the Marxist revolt indeed is a revolt against this individualistic bias. The Marxist solution, however, is quite inadequate. Class is a completely inadequate concept to bear the weight of the enormous complexity of group and integrative relationships. The closest thing to group economics comes in the theory of international trade, but even here economists have been a little uncomfortable with the national group as such, even though the nation has, in fact, turned out to be a much stronger integrative category than class.

When we look at the economics of race, the individualistic bias of economic theory leads to an insensitivity to the problems of identity, especially group identity, which are so prominent now. We are going through an agonizing struggle to find the proper place of racial groups within the integrative structure. The liberal philosophy of integration owes a great deal to economics and the implicit assumption of economics that the individual is an individual no matter what his color, creed, or national origin. This philosophy is now running into severe criticism because of its apparent denial of the reality of groups. We face a par-

adox here that, whereas the individual person is the only ultimate social reality, it is a reality which is sharply conditioned by the groups within which the person has grown up. The individual, furthermore, cannot even be an individual unless he has some sort of group identity. Just which group identities are the most fruitful, however, must remain an open question and we certainly cannot assume that a common skin color or a common anything automatically creates community and group identity.

Oddly enough, our statistical information system has a much stronger "racial" structure than does economic theory. In economic theory there is practically no recognition that race exists at all as social phenomenon. In our statistics, however, especially in the census, we categorize people by race quite without regard as to whether this is, in fact, the most significant and useful category. We have a lot of information about the subset of black people, but we have virtually no information about the subset of red-headed people or left-handed people, or diabetic people, or such other group that might have some significance in society. In some ways this has given the racial groups an artificial status in the statistical information set which could easily lead to a failure to obtain information about other significant groups. In different ways, neither economics nor statistics has recognized adequately that the problem of what groups are significant in society is both very important and not to be taken for granted and requires a great deal of research, most of which we have not done.

Whatever groups we decide are significant in society, it is clear that economists constitute one of them. There seems to be an inescapable dilemma here that intellectual products like economics are produced by intellectuals and intellectuals tend either to be middle class or to become so. The only remedy for the kind of culture-boundness which this produces is awareness. Culture-boundness is most dangerous when it is unconscious, as it frequently is, and when it consists of taking things for granted simply because one has grown up with them and has never questioned them. The very fact that the sciences are questioning subcultures tends to release them in some degree from the culture-boundness of traditional cultures in which no questions may be asked. Science, however, is not a complete culture. A scientist, once he emerges from his laboratory, his classroom, or his library, becomes an American or a Russian, an Adventist or a Communist, and he is highly likely to be middle class and to participate in his private life in middle-class habits and ways

of life. The Temperate Zone bias of science we have already noticed. Its class bias is perhaps less noticeable, but under some circumstances it may not be negligible. Social sciences, of course, have to be particularly aware of this insofar as they are dealing as scientists with the very things that they tend to take for granted as human beings and citizens.

In a complex society, everyone has to relate to more than one culture and the economist is no exception. Up to a point, economics is a remarkably homogeneous subculture. Even socialist and capitalist economists share an enormous amount of common ground, in spite of the different ideologies and values. Similarly, Catholic economics is not very different from secular economics, or black economics from white economics. Nevertheless, the larger culture within which the economist

finds himself inevitably has some influence on his attitudes and his priorities within his own profession. The economist who is teaching economics, say, to trade unionists or to Catholics or to socialists or to black people or to Indians would have to be unusually insensitive not to be aware of the culture with which he is intersecting. Oddly enough, the one cultural minority to which, as teachers, we make the fewest concessions is the great cultural minority of youth, which increasingly is developing a subculture and a language, a set of values and interests of its own, different in many respects from the rest of society. Perhaps then the greatest source of culture-boundness among economists is being over thirty, a condition for which there seems to be no cure, but in which, like all others, self-awareness may help.

RECOMMENDATIONS FROM NASHVILLE CONFERENCE ON ECONOMIC CURRICULUM IN BLACK COLLEGES*

By FLOURNOY COLES
Fisk University

Purpose and Objectives

The Conference was designed to ascertain the extent and quality of the teaching of economics in the predominantly black colleges and universities in the United States. Toward this end, the Conference addressed itself to the four major areas—i.e., curricula, teaching methodologies and techniques, textbooks and library availabilities, and research—involved in the effective teaching of economics. Built into the Conference structure was a mechanism for the making of recommendations for improvement in each of these four essential areas.

Participants

Representatives from thirty-four predominantly black colleges and universities attended the Conference. All representatives were engaged in the teaching of economics courses at their respective institutions at the time the Conference was held, and approximately 50 percent of them were or are now heads of the departments in which economics courses are taught in these institutions. In addition, five consultants were in attendance. These were Professors G. L. Bach (Stanford University), K. E. Boulding (University of Colorado), Rendigs Fels (Vanderbilt University), Phillip Saunders (Carnegie-Mellon University), and Harold Williamson (Northwestern University). Four other economists participated as observers and discussants: Professors Herman Freudenberger (Tulane University), Donald Harris (University of Wisconsin), Thomas Palm (Portland State College), and Harold Vatter (Portland State College).

Problems

At the very beginning of the Conference—at the keynote banquet speech given by Dr. Phyllis Wallace, Director of Research, Metropolitan Applied Research Center, Inc., New York—major problems began to emerge. Focusing on the need to increase the supply of black economists and on the economic consequences of racial discrimination, Dr. Wallace: (1) stressed the need to pro-

duce technically competent and intellectually versatile blacks who may not share the dominant values of the white upper-middle class and who are willing to work for change from within the system; (2) noted that the social science community, and especially economists, have become highly specialized, introverted, and lacking in feeling and vision; (3) observed that black students in black and white colleges and universities have forced a reappraisal of the "rules of the game" by administrators and faculty; (4) detailed the difficulties in teaching economics at predominantly black schools (e.g., isolation from the real world of economics, abandonment by the profession, heavy teaching loads, limited funds for books and teaching aids, etc.); (5) deplored the apparent lack of interest on the part of the profession as a whole in economic problems of direct concern to the black community; and (6) analyzed at considerable length the socioeconomic impact of racial discrimination within the profession.

1. *Curricula.* The major problems in this area were outlined in a paper presented by Dr. R. Grann Lloyd, Chairman of the Department of Economics and Business at the Tennessee A. and I. State University. These were: (1) 44.9 percent of the predominantly black colleges and universities offer less than ten courses in economics, 36.7 percent offer ten to fifteen courses, 12.2 percent offer sixteen to twenty courses, and 4.1 percent offer twenty-one to twenty-five courses; (2) there are major weaknesses in the course offerings, programs, and organization in economics in the predominantly black schools; (3) historic limitations on practical occupational choice and employment opportunities for blacks are directly related to economics curricula inadequacies and misdirections in the black colleges and universities; and (4) black economics courses and/or programs versus the application of economic analysis and statistical techniques to the understanding of the problems of the black community.

2. *Teaching Methodologies and Techniques.* The Conference attempted to develop answers to four major questions in this area. These were: Are traditional teaching methodologies and techniques applicable without modification to black colleges and universities? If not, what modifi-

* Conference on Economic Curricula in Black Colleges, Fisk University, under Ford Foundation Sponsorship, April 17-20, 1969.

cations are needed? Are there opportunities for or benefits to be derived from programmed instruction in economics in black schools? Of what benefit would cooperative education or work-study programs between employees and black colleges be? To what extent should black communities be used as laboratories in the teaching of economics in black schools?

The attempts to find answers to these questions surfaced several problems with which the black schools are confronted. Chief among these were: (1) impediments to more effective teaching of economics represented by administrative obstacles, inadequate salaries, insufficient budgets, the lack of sabbatical leaves, excessive teaching loads, and heavy committee responsibilities; (2) the lack of experience of current faculties in programmed instruction; (3) difficulties in determining the usefulness in predominantly black schools of programmed instruction in economics; (4) the need to assure the acquisition of credit for enrollment by students in cooperative education programs; (5) the development of cooperative education programs which will not extend the length of time a student would have to remain in school; and (6) the necessity of developing and implementing programs in black communities which are not either demeaning to the communities or perceived by the residents as such.

3. *Textbooks and Library Availabilities.* Dr. Jesse Gloster, Chairman of the Department of Economics at Texas Southern University, presented a paper which detailed the major problems in this area. These included (1) insufficiencies of and inadequacies in library holdings, (2) the choice of textbooks without regard to their relevance to particular situations, and (3) although standard and thorough textbooks are being used, in many instances these are not being supplemented with more relevant materials when the opportunity to do so arises.

4. *Research in Predominantly Black Schools.* By and large, there is a paucity of economics research being conducted in the predominantly black colleges and universities, which suggests to outsiders stunted intellectual growth and development on the part of current faculty and makes more difficult the recruitment of young black economics Ph.D.'s. The reasons cited for the paucity of research were (1) poor research facilities, (2) excessive teaching loads and committee responsibilities, (3) negative attitudes towards research on the part of school administrators, (4) lack of funds, (5) the absence of a community of scholars, (6) the absence of graduate programs, and (7) the isolation of teachers in most predomi-

nantly black schools from what is going on in major research centers.

Recommendations

1. *Dr. Phyllis Wallace.*

a) Identify the numerous opportunities for undergraduate students to secure summer or interim employment, in which they have opportunities to utilize their training in economics.

b) The transportation of economics majors to Washington, D.C., for a series of lectures and briefings by black economists who have attained high positions in the federal government.

c) An evaluation of the various summer institutes for black economics majors, and a combining of the most successful elements of each into a single program for future institutes.

d) The establishment of several centers, within the black college complex, with each center specializing in courses or areas (e.g., manpower development and utilization, urban renewal, regional planning, etc.) and students sent there for one term.

e) Faculty enrichment programs.

f) A clearinghouse for data on numerous social welfare programs, to be established at one black college or university with data available to all participating members.

g) Assistance to the teachers of economics in predominantly black colleges and universities by professional societies.

h) The establishment of a first-rate economics department at one predominantly black school, to be financed and assisted by foundations, government, and the profession.

i) The employment of a full-time person to coordinate the programs designed to increase the supply of black economists.

2. *Curricula.*

a) The establishment of a minimum core curriculum at each of the predominantly black institutions, in the interest of best meeting the needs of black students, which would include six semester hours of principles, and three semester hours each of macroeconomics, microeconomics, and statistics.

b) Supplementing the core requirement of fifteen semester hours with money and banking, economic thought, labor economics, international trade, economic development, public finance, urban economics, consumer economics, labor relations, and/or courses concerned with the economic problems of the black community.

c) Institute curricula changes based on both

the preparation of students for graduate study in the field and the preparation for useful work in the black community.

d) Maintaining an appropriate balance between student preferences and professional considerations.

e) In the area of economic development, subject to further investigation that which is treated as parameters in conventional economic theory.

f) The "housing" of economics in a separate department, if possible, as opposed to combining it with business administration or other social sciences.

g) The requirement of some mathematics for all students majoring in economics, with the minimum requirement to include calculus.

h) The requirement for economics majors of courses in American history, American government, logic, sociology, English, and the life sciences.

i) The integration of materials which would go into the making of a black studies program into existing courses, wherever possible, as opposed to the creation of a separate black studies program.

j) The preparation of a suitable alternative to the principles sequence as a means of interesting and motivating students.

k) Internships or other forms of cooperative education with business and government.

3. *Teaching Methodologies and Techniques.*

a) Resort to modifications and innovations, to the maximum extent feasible, in the interest of the best preparation of students.

b) More independent student participation through individual research projects and reading assignments.

c) A more extensive use of workbooks.

d) The use of simulated games.

e) Pooling arrangements with other schools to bring in distinguished economists, and to develop and use such aids as films, transparencies, and telelectures.

f) Investigation of the Test of Economic Understanding, the Test of Understanding College Economics and/or other "normalized" tests as means of evaluating the effectiveness of programmed instruction.

g) More meaningful involvement with each other of black schools, black students, and black communities.

h) Consider examinations more realistically as integral parts of the educative process, including the use of such examinations as the Graduate Record Examination as a teaching device.

4. *Textbooks and Library Availabilities.*

a) Available textbooks should be examined closely and choices by instructors and departmental heads should be made with due regard to relevance to particular situations.

b) "Usual" economics texts should not be stocked in libraries; thus reserving scarce funds for the acquisition of reference materials.

c) Because of the many new textbooks which are coming onto the market each year, some consideration should be given to changing textbooks fairly frequently rather than sticking to up-dated versions of old textbooks.

d) In stocking libraries, care must be exercised to give priority to areas which are most relevant.

e) Cases, kits, and games should be used to supplement textbooks in the interest of greater student motivation and interest.

f) Films, transparencies, and other visual aids should be stocked in libraries to a much greater extent.

5. *Research in Predominantly Black Schools.*

a) Eliminate or reduce the practice of giving all courses every year, so that black scholars will have more time to do research.

b) A more intensive search for research funds by black schools.

c) The inclusion of black schools in the circuit of mimeographed and other papers circulated by large research centers for comment.

d) The development of clearinghouses within the black college complex for research papers, as suggested by Dr. Wallace.

e) Encouragement and assistance in three areas of research by black economics scholars—policy-oriented research, research in the legitimate interest-areas of black scholars, and research in the interest of other scholars in black problems.

f) Faculty-student research in community problems.

6. *Others.*

a) The establishment of a Permanent Advisory Group to assist the black colleges and universities. The group would consist of: G. L. Bach, Stanford University; K. E. Boulding, University of Colorado; Andrew Brimmer, Federal Reserve Board; Rendigs Fels, Vanderbilt University; Walter Heller, University of Minnesota; Vivian Henderson, Clark College; Sir Arthur Lewis, Princeton University; Emmett Rice, International Bank for Reconstruction and Development; Phillip Saunders, Carnegie-Mellon University; Henry

Villard, City University of New York; Harold Williamson, Northwestern University.

b) A subsequent conference during the 1969-70 academic year, to include a mixture of schools represented and schools not represented this year, with at least one student majoring in economics from each school represented in attendance.

c) Specific provisions for a discussion of "Black Studies" during the subsequent conference.

d) The immediate development of a program of action in all areas of concern at the Conference, which would involve the Permanent Advisory Group working with a subcommittee or subcommittees established from among those attending the Conference. The overall mandate of this combined group would be to design a comprehensive plan to produce more black economists.

e) A round table meeting to be arranged for the next meeting of the American Economic Association, in New York in late December, 1969, for the purpose of bringing to the attention of A.E.A. members the problems of economics instruction and research in the predominantly black schools. The persons suggested to make the necessary ar-

rangements are: Dr. Phyllis Wallace (Chairman), President Vivian Henderson (Clark College), Professor Cleveland Chandler (Morgan State), and Professor Flournoy Coles (Fisk University).

f) Direct assistance to the black departments of economics by the economics profession (e.g., the assignment for a year of Dr. Paul McCracken at Talladega College, recently). The costs of such assistance, most probably, could not be assumed by the schools needing such assistance.

g) The establishment of a fund to be used to send black economics faculty members to professional meetings.

h) The sending of up to twenty-five carefully selected black students to the Economics Institute for Foreign Students at Boulder, Colorado, each summer.

i) The development of a test to replace the Graduate Record Examination to be used by graduate departments of economics.

j) The development of means whereby the black schools can compete more effectively for black economists.

k) The establishment of summer institutes for black students on black campuses.

AN AFFIRMATIVE ACTION PLAN FOR THE ECONOMICS PROFESSION*

By CLEVELAND A. CHANDLER

Morgan State College

I. Assumption and Values

The proposed plan is based on the following assumptions:

1. That the relative distributions of aspirations and innate abilities among whites and blacks for pursuing the economics profession are similar if not identical, given comparable exposure to training and access to employment opportunities.

2. That the shape of the performance distribution curve of whites and of blacks who would enter the economics profession is similar or identical, given (1).

3. That the quality and the quantity of output of the economics profession would be maximized and optimized at comparable levels of relative utilization of blacks and whites in the various areas of specialization, given (1) and (2) above.

4. That balanced utilization of blacks and whites in the economics profession at optimal output levels is preferred to imbalanced utilization, misallocation of human resources, nonoptimal levels of output, and maldistribution of income.¹

5. That the nature of the product to be produced by blacks and whites in the economics profession is essentially the same, notwithstanding variations in emphasis and applications stemming from particular needs, orientations, commitments, and historical experiences of blacks and whites.

II. Some Elements of a Model

The first part of Figure 1 is given by nature, and in the absence of clear empirical evidence to the contrary can be questioned only on the basis of bigotry. The second part of Figure 1 may differ for whites and blacks only as a result of differences in training exposure and opportunity access in the past and in the present. If blacks are distributed with respect to aspirations as shown by the broken curve, it is the results of differ-

ences in black-white training exposure and perceived opportunity access. An effective plan must remove the distortion from the black or the white aspiration distribution that is due to contrived and inadvertent differences in training exposure and perceived opportunity access. The broken curves for blacks must be shifted to the right and the skewness in it must be eliminated, if the shape of the aggregate curve is normal under total equality.

Under the conditions of equal training exposure and equal opportunity access, populations of black and white economists are distributed according to a single set of performance standards. However, actual performance standards are not identical for the two populations. Dominant white standards place high priority on research and publication, especially in the academic branch of the profession. Black economists hardly participate in determining this set of performance standards, and usually must compete from positions of adverse working conditions and disadvantage. In reality, black economists perform by another set of standards that place a higher priority on teaching and the amelioration of severe practical economic problems in the black community. Therefore, they perform according to two sets of standards, which are not mutually exclusive, but they are distinctly different in some respects. In relating to double sets of standards, the performance of black economists is depressed in both lines of activity. Or, in relating to one set, the performance of black economists is distorted. And the performance of white economists is inflated and distorted by a single set of standards which does not include some of the criteria and priorities of a comprehensive set of performance standards. An effective plan must consolidate performance standards for white and black economists so that the professional achievements of all economists are a true reflection of innate abilities and aspirations of white and black economists under equal training exposure and opportunity access.

If white innate economic talent per capita is more fully utilized than black innate economic talent per capita, then there exists black talent yet undiscovered and underdeveloped. The utilization of this talent yields a high ratio of outstanding economists, which shifts the total output curve of the economic profession upward disproportionate-

*I am indebted to Vernon Dixon, of Princeton Univ., and colleagues at Morgan for critiques of the original draft of this paper that resulted in useful revisions. Remaining mistakes are mine.

¹In the following discussion, I do not consider the cost of implementing the proposed plan. I assume that the costs are not prohibitive; that a democratic society is committed, in fact, to full manpower development and utilization as well as to equal opportunity; and that the social benefits far outweigh the social costs.

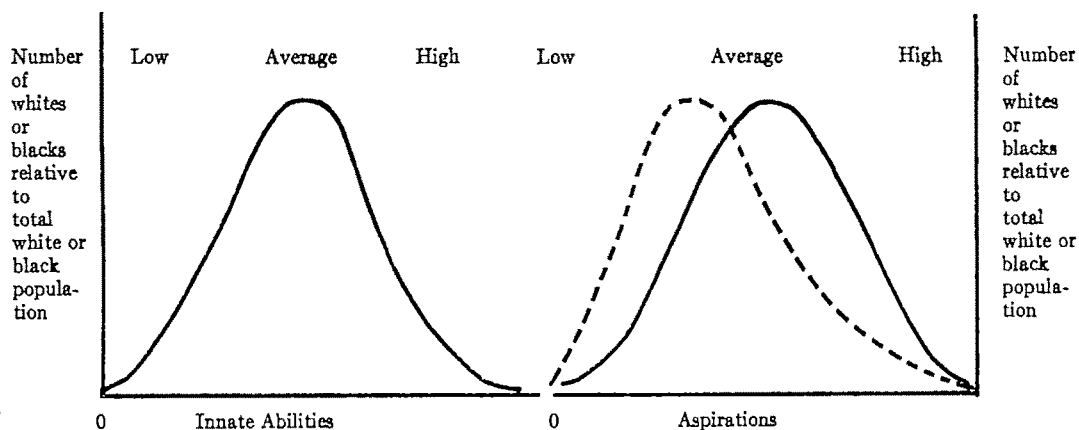


FIGURE 1
RELATIVE DISTRIBUTION OF INNATE ABILITIES AND ASPIRATIONS IN WHITE
AND BLACK POPULATIONS

ly. These additional economists, through increasing the scale of inputs, create scale economies for the profession. The discovery, development, and utilization of gifted black talent shift the point of optimum output further upward and outward on the economics profession's total product curve. Some improvement in the quality of professional output could result from cross-fertilization of racial balance in the various fields of specialization. All these possible effects bring about greater profes-

sional status and efficiency for economics as a discipline.

Thus, a proposed plan must reduce the pool of undiscovered, underdeveloped, and underutilized economic talent in the black population; thereby improving and increasing the output of the economics profession and enhancing its professional status, effectiveness, and respectability.

We postulate one body of economic thought, theories, and principles whose policy implications

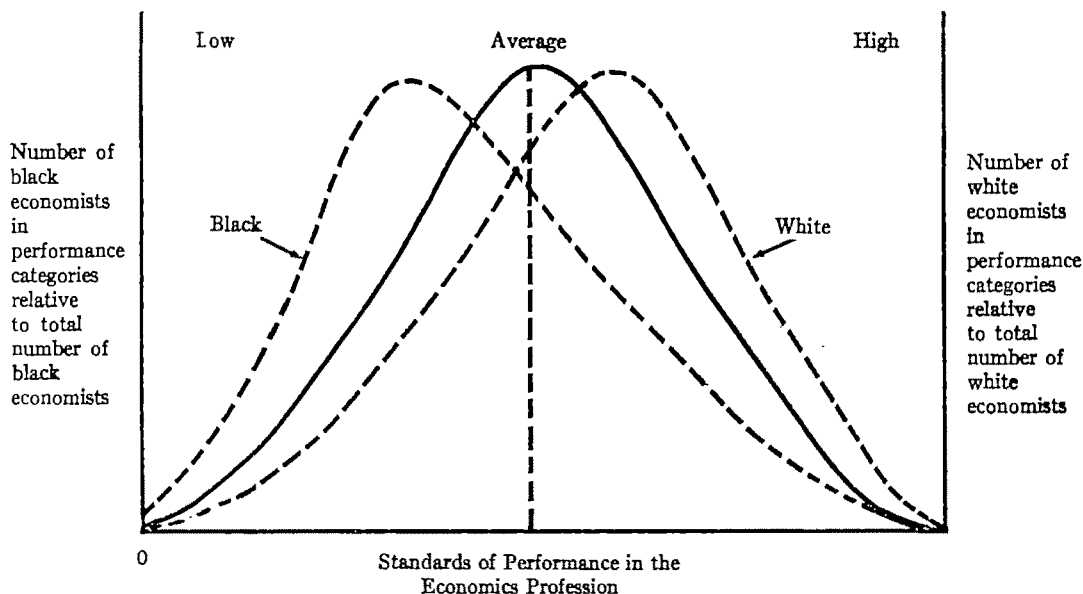


FIGURE 2
RELATIVE DISTRIBUTION OF BLACK AND WHITE ECONOMISTS
ACCORDING TO WHITE PERFORMANCE STANDARDS

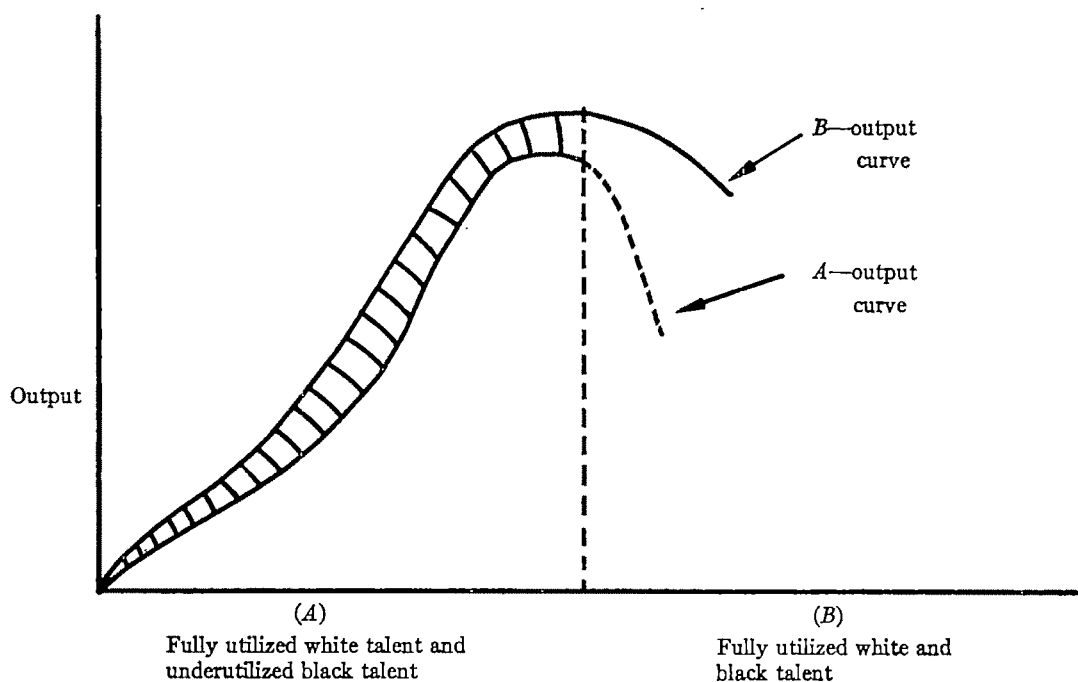


FIGURE 3.
HYPOTHETICAL OUTPUT CURVES FOR UNDERUTILIZED AND FULLY UTILIZED
BLACK TALENT IN THE ECONOMICS PROFESSION

and applicability for practical economic problems refer to the total human population, the white component, the racially integrated component, and the black component. Some theories and principles as presently formulated may be more applicable and relevant to one segment of the total society than to others. Undoubtedly, opportunities exist for ingenuity and creativity in the modification of particular economic principles, theories, and models to increase their relevance and applicability. New economic concepts, theories, and principles may emerge from the study of economic conditions in the black community. Some, if not all, of the newly discovered theories and modified economic relations may be generalized to other components of the total human population and the entire society. Since one global body of economic thought is the frame of reference for the professional economist, this is the body of knowledge to which I refer in setting forth my proposals to increase the supply of black economists.

An effective plan must improve the relevance and the applicability of the science of economics to economic phenomena in diverse human populations. In testing the relevance and applicability of

existing economic theory and principles to particular human populations, a more generalized body of economic knowledge is envisioned in which economists from different ethnic backgrounds can be trained and/or can practice in different human populations.

Thus, if white-black differences in training exposure and opportunity access are removed, if performance standards for white and black economists are consolidated and equalized, and if the relevance and applicability of the science of economics in regard to divergent human populations are improved—and by so doing, the body of economic knowledge is enriched—then the status and the state of the economics profession will be enhanced. Hence, the proposed plan also coordinates the several strategies for increasing the supply of black economists so that they constitute an affirmative action plan (program) for the economics profession.

III. Operational Aspects of the Problem

According to current estimates, there are about 25,000 economists in the United States. In 1966, a total of 13,150 were registered with the National Science Foundation, among 242,763 regis-

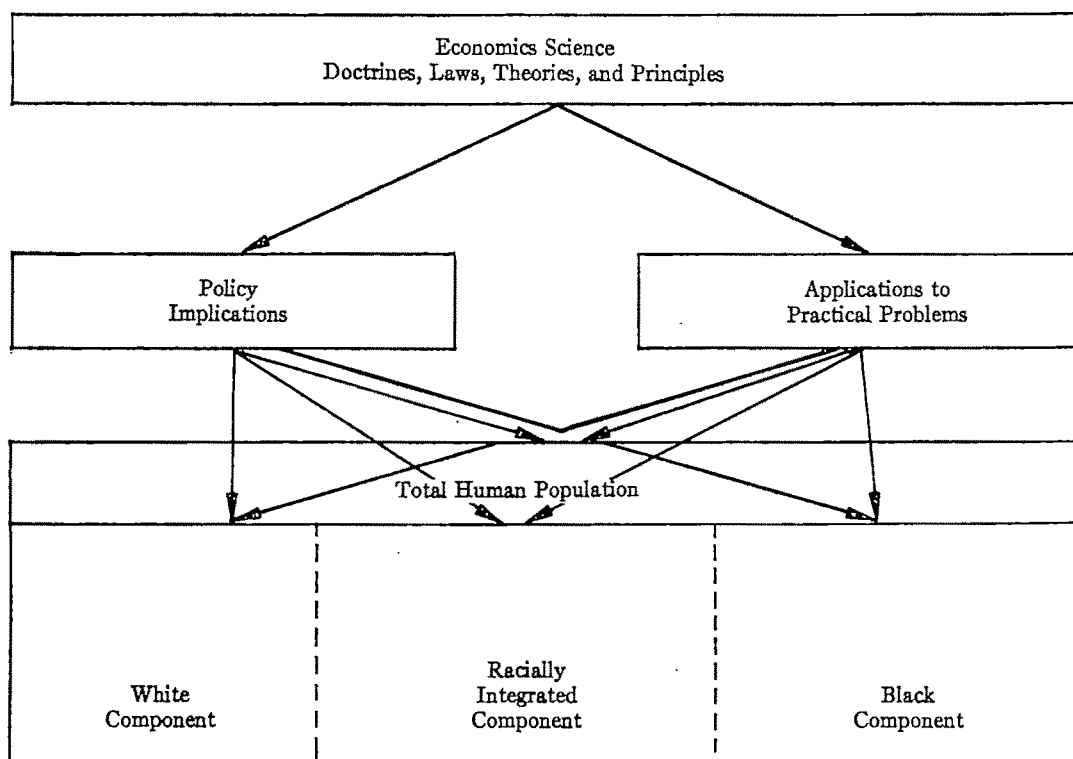


FIGURE 4.
DIAGRAM OF ECONOMICS SCIENCE RELATIVE TO
POLICY IMPLICATIONS AND APPLICATIONS

tered scientists.³ Between 1964 and 1966, the number of registered economists increased from 12,000 to 13,150. Apparently about one-half of all U. S. economists are included in the NSF National Register of Scientific and Technical Personnel, because the 1960 Census reported 22,000 economists. In any event, the total supply of economists is growing, presumably in response to a faster growing demand for economists. This sug-

³*National Register of Scientific and Technical Personnel* (National Science Foundation, Washington, D.C.), "Summary of American Science Manpower, 1966"; also, *American Science Manpower, 1966*. According to the summary, the National Science Foundation is required by the National Science Foundation Act of 1950 to maintain a register of scientific and technical personnel and in other ways provide a central clearinghouse for information covering all scientific and technical personnel in the United States. Since 1953, the National Register of Scientific and Technical Personnel has been maintained as a cooperative endeavor of the National Science Foundation and the scientific community, as represented by the scientific professional societies. In the field of economics the survey for the National Register is conducted by the American Economic Association.

gests that recommendations for increasing the supply of black economists can be adopted and implemented with little or no threat to the job security of current and projected stock of economists.

The 25,000 U. S. economists are equivalent to one economist per 8,160 persons in the total U. S. population. On a pro rata basis, there would be between 2,500 and 3,000 black economists in the United States now. According to available data, the actual number of black economists is far short of a proportional share of the total.

One survey conducted by Richard R. America, with the Stanford Research Institute in 1968, revealed a total of only 70 black economists. At that time, 32 of these economists held teaching positions at predominantly black colleges and universities, 18 held positions at predominantly white colleges and universities, at least 11 were employed by the federal government, and the remaining 9 were employed by private corporations and research organizations. Although this survey did not claim coverage of all black economists in the United States, it seems fairly certain that the total

does not exceed several hundred.³ Therefore, at least a tenfold increase in the number of black economists is required to eliminate the deficit.

Our plan for increasing the supply of black economists is concerned with reducing imperfections in the labor market as such rather than with a shift in the supply curve with respect to price. Given the salary structure for economists, the plan is to specify actions that reduce excessive impediments to the flow of black economists into the preparatory pipelines of the economic profession. Concomitantly, an increased number (stock) of black economists is to be distributed more widely in the different facets and at different levels of the economics profession. Mean-time, economic theory and practice, the output of the profession, must be enriched by these changes. Some of the components of a plan are presented below.

IV. An Affirmative Action Plan

The Affirmative Action Plan is delineated, partly, in terms of a time frame, operational mechanisms, and expected results for each strategy and related actions. Hence, the plan consists of five series of actions to be completed on schedule, in a prescribed manner, by a target date. The recommended target date is 1975.

The implementation of Strategy (1) and its related actions implies more than an increased supply of black economists. In addition to the enhancement of the posture of the economics profession, this also constitutes a model of human resource development of a small segment of the total labor supply, which penetrates the entire labor market mechanism. Economic education below the collegiate level would be improved vastly.⁴ Labor market information on opportunities and qualifications of economists would be distributed widely in the schools and in training and career counseling channels. Economic offerings at predominantly black colleges and universities would be funded at levels necessary to transform devel-

oping areas into fully developed and advanced areas. This would, then, attract increased numbers of better prepared white students and professors to predominantly black colleges and universities as better prepared black students and professors are attracted to predominantly white colleges and universities.

The National Science Foundation and the American Economic Association could institute a comparative study of black and white registered economists. Moreover, such a study could be utilized to extend the coverage of the Register, especially among black economists, since it is likely that a low proportion of the relatively few black economists are now included. Undoubtedly, analyses of these data would be helpful for planning, implementing, and evaluating a program for increasing the utilization and participation of black economists in the various phases and levels of the economics profession. Increased volume supply of black economists with wider distribution would give the labor market mechanism, under conditions of equal opportunity, a chance to work.

Fortunately, the strategies and related actions listed above can be instituted in the context of an expanding profession in an expanding volume of aggregate employment. Moreover, the structure of employment and salaries in the economics profession is tending more and more in the direction of a closer relationship with academic qualifications obtained through formal training rather than through experience and mere designation. A few of the findings of thorough study of the structure of employment and salaries in the economic profession substantiate these assertions.⁵

Furthermore, the salary structure of the economics profession compares favorably with the salary structures of the twelve or so professions included in the National Register.⁶

In 1964, the 10,000 economists who reported full-time professional work to the National Register earned salaries ranging from less than \$5,000 to more than \$100,000. The arithmetic mean salary for these economists was \$13,670 and the median was \$12,000. Eleven percent of these economists, the largest concentration, fell into a salary range of \$9,500 and \$10,500. This suggests that there may be considerable leverage for expansion of the salary structure from the bottom,

³ A list of 126 black economists was distributed in the Black Caucus at the 1969 Annual Meeting of the American Economic Association. Sixty-eight, or 54 percent, of these economists held terminal degrees in economics at that time and 37, or 54 percent, of the 68 economists with terminal degrees held positions in predominantly black colleges and universities.

⁴ See National Task Force on Economic Education; *Economic Education in the Schools Report* (Sept., 1961). Materials Evaluation Committee, *Report* (Oct., 1961); Special Textbook Study Committee of the Committee on Economic Education of the American Economic Association, *Economics in the Schools* (A.E.R., Mar., 1963, Pt. 2).

⁵ Committee on the NSF report on the profession (N. Arnold Tolles, Alice Hanson Jones, and Ewan Clague), *The Structure of Economists' Employment and Salaries, 1964* (A.E.R., Dec., 1965, Pt. 2).

⁶ Economics, statistics, psychology, sociology, linguistics, physics, chemistry, mathematics, biology, meteorology, earth sciences and agriculture.

SOME ELEMENTS OF AN AFFIRMATIVE ACTION PLAN FOR INCREASING THE SUPPLY OF BLACK ECONOMISTS

Strategy and Related Action	Time Table	Operational Mechanism	Expected Results
Strategy (1): Equalization of training exposure and opportunity access	1970-75, on demonstrated and intensified basis; and continue forever, unaccentuated	The labor market and labor market feeder system; labor market information systems and economic education in secondary schools	Increased number of actual and aspiring black economists
Action (1): Intensify economic education effort in elementary and in secondary schools, especially those with predominantly black enrollments	Phase in 1970-72; stabilize at adequate level, 1973-75; continue at adequate level indefinitely and expand to general population	Committee on Economic Education; A.E.A.; Joint Council on Economic Education; Committee for Economic Development	Increased economic literacy in unreached component of pupil population with increased flow of black pupils in the preparatory pipelines of the economics profession
Action (2): Develop and disseminate literature and related information in secondary pupil populations, which clearly identify the economics profession as a career alternative, with statements of professional opportunities and requirements; also disseminate in out-of-school population	Phase in, 1970-72; stabilize at adequate level, 1973-75; continue at adequate level indefinitely in general population	Same as in Action (1) above, and secondary schools, the federal-state employment service system, and other employment counseling and career guidance-type organizations	Increased numbers of college and university students with majors and minors in economics whose training and occupational choices are more consistent with their native talents, interests, and aspirations
Action (3): Upgrade and expand the economics curriculum along with faculty, facilities, and teaching-learning technology in predominantly black colleges and universities	Phase into saturated level, 1970-72; stabilize at intensified level, 1973-75; continue and diffuse among all colleges and universities, indefinitely	Continuation of Ford Foundation, Fisk University Conference on Economic Curriculum and IBM-UNCF Summer Institute for Teachers of Economics, with action content	Increased numbers of strong black graduate students of economics and increased supply of blacks professionally trained for the different areas of specialization in economics
Strategy (2): Unify performance standards, and related actions	1970-75, and forever	Governing boards of colleges and universities, federation of accreditation associations; generalized standards and procedures of examining boards and committees; validation of standardized tests in black component of population	Uniform standards of accreditation; uniform qualifying standards and professional ethics with quality control
Strategy (3): Reduce undiscovered, underdeveloped, and underutilized talent for the economics profession in the black population, and related actions	Reduce to comparable level in the white population 1970-75 and minimize in the general population forever	High school counselors through their associations; more contacts between college and university students and economists, and indirect contacts through collegiate counselors and their associations	Increased numbers of better qualified black students of economics and professionals in branches of economics for which they are best suited
Strategy (4): Improve relevance and applicability of economics science for various facets of economic phenomena in divergent human populations and related actions	1970-72, intensified application of economics science to black ghettoized populations in urban centers; 1973-75 extend to other population concentrations	American Economic Association, Southern Economic Association, National League of Cities, selected government agencies; various minority economic organizations	Diffusion of economic knowledge and expertise to populations with chronic economic problems; generalization and validation of economic theories and principles in these populations
Strategy (5): Coordinate and concert all strategies and related actions as a comprehensive and massive program for increasing the supply of black economists	1970-72, build program to peak level in black population; 1973-75, stabilize and routinize program, and extend to other components of the general population	American Economic Association in cooperation with other economic associations, societies, foundations, and organizations	Marked increase in total number of black economists in various areas and fields of specialization, and greater stature and integrity for economists and their profession

since, apparently, there is an acute shortage toward the top of the highly skewed distribution of economists by salary.

The median salary of economists in 1964 was the same as that of the physicists and the statisticians and was from \$1,000 to \$3,000 above those of the other nine professions included in the National Register. Again, this suggests an opportunity to increase the supply of black economists by the attraction of a favorable comparative salary structure in the context of an expanding labor market situation. Increased opportunity access

and training exposure for blacks could be provided with minimal or even negligible alteration of the opportunity spectrum for others. Undoubtedly, a substantial number of blacks could be absorbed by the shortage that already exists in the profession.

Another point in this regard is that the part of the economics profession most immediate to the preparation and training of an increased supply of black economists, the academic component, is the least likely candidate for the absorption of the increased supply. Therefore, protective behav-

ior on the part of this component of the profession is unjustified.

Due to past and present labor market imperfections, there is already a backlog of black economists in the academic end of the profession who would move to other segments with an adequate supply of replacements and with moderate mobility inducements. Second, salaries in other components of the economics profession are higher. For example, the committee on the NSF report on the profession also found that industry or business, where one-third of the registered economists were employed in 1964, paid the highest median salary of \$14,400, exceeding that of any other of the eleven professions by amounts ranging from \$300 to \$5,400. The federal government, where one-tenth of the registered economists worked, paid the next highest median salary of \$13,200. Forty-two percent of the registered economists were employed by educational institutions with a median salary of \$10,100, which was lower by \$100 to \$400 than those paid to professions of agriculture, statistics, and meteorology, but from \$100 to \$1,400 higher than the median salaries paid by educational institutions to the other eight professions.

In terms of salary considerations, the economics profession occupies a favorable position in the academic realm, but the salary position of academic economists relative to the salary positions of economists in other segments of the profession bear a poor relationship. Hence, we can assume that most of the newly trained economists will be least competitive, if at all, with their trainers. Moreover, the influx of newly trained economists in other segments of the labor market will not reduce but will increase the opportunities of academic economists and enhance the status of the economics profession.

Most of the analyses and proposals so far have been aimed in the direction of an increased supply of black economists with increased racial integration in training, in employment, and in the American society, generally. However, as long as white racist institutions and practices prevail in the real world, black economists cannot afford to be entirely color blind. An increased supply of black economists will be called upon to supply increased amounts of economic understanding and intelligence in the black community, where acute economic problems have been neglected.

In the contemporary racial climate a large proportion of the economic intelligence that is destined for the black community is going to be transmitted by black economists. Hence, a part of the affirmative action plan to increase the supply of

economists must, at the same instance, provide for an increased number of carriers of economics science to the black community. By saturating the black population with relevant economic science and intelligence, the base of aspiring black economists will be augmented, and an increased supply of black economists could assume varying degrees of commitment to the black community and the general community.

The American Economic Association, in collaboration with other professional economic societies, is urged to commit itself to a program that will yield from 2,500 to 3,000 black economists in the next five years. Perhaps half of these prospective economists will complete their training up to the master's level in predominantly black colleges and universities, and a much larger percentage will graduate from predominantly black secondary schools. Above the master's level, almost all economic education is provided by predominantly white universities.

At this point a dilemma and a challenge emerge. The dilemma is faced by the prospective black economists who must adjust, progressively, to white-oriented and controlled work situations and at the same time act more forthrightly on their commitment to make economics more relevant to problems in the black community. The challenge is faced by the economics profession. It must equalize the quality and the scope of training for prospective black economists in predominantly black and white schools, upgrade and expand economic education in general, and increase the applicability and relevance of economics in the total human population, including the black component, the white component, and the racially integrated component.

V. Summary and Conclusions

An effective plan to increase the supply of black economists is based on certain assumptions and values regarding the distribution of innate abilities and aspirations within component groups in the total population. Granted these assumptions, gross disparity in the relative supply of black economists must be attributed to black-white differences in training exposure and perceived opportunity access.

Also, an effective action plan, under basic conditions of racial integration and factual equal opportunity, must unify performance standards in training and employment. These requirements can be met only through substantial change in labor market institutions governing the development and utilization of human resources in the economics profession.

The dilemma of two sets of performance standards and dual sets of commitments faced by black economists must be resolved, either in the direction of a racially integrated system of economic education in a racially integrated society or in the direction of dual societies and systems of

economic education, with parity of training exposure and opportunity access. In a large measure, the challenge afforded by these options and the responsibility to move affirmatively in one direction or the other rest with the decision-makers in the economics profession.

THE DI-UNITAL APPROACH TO "BLACK ECONOMICS"*

By VERNON J. DIXON
Princeton University

I. Introduction

Economics is Economics and there is no "Black Economics"! This position implies that economics is a discipline with a received body of knowledge, analyses, methodology, and attributes that are "universally" applicable. In this context, the laws of economic behavior or economic concepts are considered race free and color blind. The purpose of this paper is to provide some support for an alternate proposition: economics is economics and there is also black economics. Furthermore the analysis presented here may provide some incentive for American black students to enter the profession as well as suggest, to both black and white economists, a new approach to the formulation of economic theory.

Before undertaking this task, it will be helpful if I first examine some comments Professor Boulding made in his 1968 Presidential Address to this Association [2, pp. 1-4]. He points out that economics as a science cannot be divorced from the effects of culture; i.e., from the effects of a set of values common to a group. In other words, economics is formulated by people who, in turn, are shaped by their environment. Consequently, economics will in some sense reflect culture. It follows, therefore, that if there is an Afro-American culture different from white American culture, then the shared values unique to the black American group can lead to the development of black economics. Accordingly, black economics refers to a body of theory that is consistent with or based on the cultural uniqueness of Afro-Americans.¹ This cultural input approach to

economics seems to raise the specter of a deluge of special modes of analyses—many cultures, many economics—which is in sharp contrast to the notion of a single "universal" economics. What, then, is the relationship between economics and culture?

This question, moreover, is related to an area not usually considered the province of economists: the process of human learning. As Boulding indicates, economists have contributed more to a description of what is learned rather than to a description of the genetics of knowledge. For example, he points out that the preference functions embody what is learned about values. Since most human preferences are learned, Boulding, citing Veblen, notes that tastes cannot be assumed as given in dynamic theory. Rather, dynamics must speak about the process by which preferences are learned. What, then, is the relationship among economic theory, the process of learning, and culture? In providing an answer to this question, I shall set forth the characteristics of two alternative processes of human learning, namely, the either/or approach and the di-unital approach, and discuss how they affect one's view of culture and economics.² I shall then use this general analysis to indicate a path along which future economic theorizing might proceed.

II. Conceptual Frameworks, Economics, and Culture

In this paper, I define human learning as a process by which an individual or group inculcates concepts, values, customs, views, etc., resulting in a way of perceiving or evaluating data, communicating ideas, and regulating behavior. The focus here is on one characteristic of this process, for any given process indicates a particular way of forming ideas, a particular way of ordering experience, a particular way of knowing. In other

plained within a coherent, structured, distinct, Negro culture which represents a synthesis of African culture in contact with American European culture under slavery to the present day" [1].

²I first set forth the technical rudiments of the analysis of these learning processes in [4] while the application of this analysis to the cultural values of black and white America appears in [3]. A more extensive paper using this analysis and a related theory of racial conflict is currently being prepared for publication.

*I am indebted to several people at Princeton who commented on earlier versions of this paper. Professor W. J. Baumol helped immensely with his comments on exposition and the scope of the analysis. Marc Gaudry made valuable suggestions in the economics area, while Aggrey Brown suggested some analytical refinements in the noneconomics area. Finally, Jane Coppock was indispensable as an intellectual whetstone in the development of my basic ideas. Needless to say, I am fully responsible for the remaining inadequacies.

¹Research confirming the existence of a distinct black culture began with Herskovits and is currently being forwarded by the linguistic and anthropological studies of Stewart (1969), Szved (1969), Abrahams (1967), Hannerz (1969), and others. Their basic proposition is that "the behavior of Negroes is not pathological but rather can be ex-

words, my concern is with identification of the main characteristic of a conceptual framework which underlies the manner in which one views culture and economics.

The proposition that economics is economics and there is no black economics is consistent with a process of human learning characterized by what I call an either/or approach. People who employ this conceptual framework order their experiences in terms of polar categories or opposites. Everything falls either into this category or into that category, but not both categories at the same time. Something is either economics or not economics. Since philosophy, anthropology, psychology, sociology, *et al.*, fall into the category of noneconomics, they are considered the province of noneconomists. In this context, culture, the concern of noneconomics, can be considered as given rather than as determinant in economic theorizing. As a consequence, economics appears culture-free.

In the area of American race relations, black falls into the category of nonwhite and white falls into the category of nonblack. Therefore the either/or habit of mind allows one to conceive easily of black and white as mutually antagonistic opposites. Analogous to the two-person, zero-sum game, any gain for blacks becomes a loss for whites and conversely. Many white Americans reflect this approach when they deal with the presence of the black man by banishing him and his culture from their own minds—the “invisible man” syndrome. In effect, the either/or conceptual framework does not permit one to recognize that the cultural uniqueness of Afro-Americans is copresent with the cultural uniqueness of white Americans. Accordingly, white and black Americans who adopt this perspective argue quite logically that there is no black economics. They simply can’t conceive of any uniquely black American values upon which economic concepts may be based.³

Economics is economics! The universality implied in this statement is consistent with the either/or assumption of culture as given. If, however, an economist views culture as determinant, he is faced immediately with the fact that there are many different cultures. This implies that there could be as many economics as there are cultures. To achieve consistency between the de-

sired universality of economics and the fact of cultural diversity, the either/or-oriented person stresses the presence of a unified world culture to the exclusion of the cultural uniqueness of various groups.⁴

In sum, if the process of human learning is either/or in character, the bipolar categorizing permits one to conceive of economics as culture-free in two ways: (1) Culture is considered as given and it is not an input into economic theorizing since culture is the concern of noneconomic disciplines. (2) Culture is a factor in theorizing, but it can be held constant, since there is only a universal culture. Essentially, the either/or mentality transforms many economics into one economics and many cultures into one culture. In this sense, the either/or conceptual framework is symbolized by our national motto, *e pluribus unum*, one out of many.

This either/or approach, however, is inadequate when one considers the following: while people, individually and collectively, have a degree of commonality or sameness, they also have a degree of uniqueness or difference. Black Americans as a group possess cultural uniqueness and at the same time have cultural commonality with white Americans. In contrast to either/or thinking, no one individual or group is all commonality and no one individual or group is all uniqueness. Accordingly, there is a universal culture and at the same time many different cultures; there is a universal economics and at the same time many different economics. In this sense an alternate national motto is, “One and yet many.”

The statement that economics is economics and there is also black economics implies an alternate conceptual framework for which I constructed a new word, “di-unital.” Webster tells us that “di” means “akin to two” or “apart.” “Unital,” the adjectival form of the word unit, refers to a “single thing that constitutes an undivided whole.” Di-unital, therefore, means literally something apart and united at the same time or something simultaneously divided and undivided—a union of opposites without inherent antagonism. Essentially, the di-unital conceptual framework enables one to see how there can be diversity, for example, among disciplines and cultures and how, despite these differences, there can be unity present in this diversity.

The main characteristic of this alternative way

³ Some either/or black Americans recognize only the presence of black culture to the exclusion of white culture. To them the development of black economics will occur after a separate black state and black economy are achieved. A call for black economics as part of the current thrust for curriculum reform seems empty for them at this time.

⁴ This unified world culture is viewed in some instances as consisting of that which is common to all cultures or all people or as an amalgamation of existing and past cultures which has been developed through improved techniques of communication and transportation.

of knowing is simply this: something is both this element and not this element at the same time. Economics is simultaneously economics and noneconomics. For example, if preference functions embody what is learned about values (economics), then to a person thinking di-unitally, they also embody philosophical notions, sociological concepts, anthropological information, etc. (noneconomics). In this sense, economics and noneconomics are copresent with each other.⁵ In addition, not only are economics and noneconomics unique and simultaneously united, but they also perpetually transform each other. Further analysis and comprehension of noneconomics enlarges not only our understanding of noneconomics itself but also concurrently informs and expands our understanding of economics. The economic concepts, for example, of Smith and Marx, while intelligible in and of themselves, become more so to the extent that we increase our comprehension of their philosophical content. Similarly, a mathematically-trained mind will raise certain economic questions not normally asked by persons with different training that lead to important insights. Yet, this proposition is just as true for economists who also possess sociological skills, philosophical skills, etc. From a di-unital perspective, therefore, the noneconomics area, although different from the concerns of pure economics, is also intrinsically linked to the latter in a manner that informs and enriches economic theorizing. As a consequence, considering culture the province of noneconomics does not sever its linkage with economics. Analyses based on independent utility functions are more likely generated by and certainly more relevant to cultures that stress the value of people as individuals than those advocating the importance of people as a group. Culture is copresent with economics; economics is not culture free.

The di-unital approach, moreover, provides some useful insight into how one may view culture itself. To me, "being black" means that I embody, without inherent contradiction and antagonism, both black culture and nonblack culture. I am at once both black and not-black.⁶

⁵The term copresent is used rather than co-existent since the latter term is usually associated with the Hegelian dialectic. While the dialectical process of thesis, antithesis, and synthesis is a union of inherently contradictory opposites, this process driven by antagonism ultimately reaches a synthesis in which only the absolute exists to the exclusion of the nonabsolute. The dialectical process ends [7]. In contrast, the di-unital process is a continual union, though a perpetually changing one of inherently nonantagonistic "opposites."

⁶Afro-Americans who think in bipolar terms rec-

Blackness, in general, refers to that particular combination of attributes (language, family structure, ethics, stock of knowledge) that makes Afro-Americans as a group culturally unique in the same way that Mexican-Americans, Jewish-Americans, Anglo-Saxon Americans, *et al.*, possess these attributes in particular mixes that set them apart as groups. For example, while black American families are more extended than nuclear, this factor, although present as well in a Puerto Rican community, interacts with the the other attributes of Afro-Americans in a unique way to produce a black culture. Different cultures therefore arise because of the complementarity of these attributes; i.e., various groups of people assign different but similar weights to these culture-determining variables, some perhaps with zero weight. In this sense, blackness is the uniqueness or nonuniversal culture component of Afro-Americans as a group.

Di-unitally speaking, there is also copresent among black people a nonunique or universal component—nonblack culture. In other words, this is the interdependent aspect of "being black." Essentially it represents the results of cultural diffusion; i.e., the interaction with the river of human civilization with its many currents of different cultures. This nonblack aspect includes not only the negative reaction formation in black people to racial persecution but also the conscious assimilation of positive contributions to human civilization from other groups in the form of scientific knowledge, political systems, technology, business organization, etc. Given a di-unital perspective, a black person views other cultures as relevant, even that of white America. He actively seeks further exposure to and knowledge of these nonblack contributions, since this serves to enrich his blackness rather than to deny, destroy, or dilute it. Yet adoption, adaptation, or rejection of these contributions still remains his choice alone. He embodies without conflict a black culture and a universal culture, the latter being those attributes shared by all cultures. Di-unitally, he is one and yet many.

Thus if the process of human learning is di-unital in character, one conceives culture and eco-

oncle the conflict between their different ethnic and national identities, their "sense of twoness," by emphasizing either their African heritage or their American heritage. In regards to the latter term, this is tantamount to accepting the position of "oneness" of either/or white Americans. To them their ethnic identity is the national identity since they are the nation's dominant group, they possess the greatest power, they create and control the institutions, they condone themselves.

nomics as copresent. In general terms, individual cultural specificity shows up in the shapes of the production and preference functions of each individual for each good and in the different weights given to the arguments of each of these functions, assuming all the arguments are present in each function. The presence of the cultural uniqueness of Afro-Americans as a group can be identified because the shape of the functions and the weights given to the arguments of the functions of each individual are similar, and different from those of individuals who belong to other cultural groups. There will be, however, some functions characterized by the fact that their shapes and the weights assigned to each of their arguments will be the same for all individuals; thus indicating the copresence of a universal culture. Consequently the di-unital "habit of mind" permits one to acknowledge the copresence of universal economics and many different economics, one of which is black economics.

III. *Conceptual Frameworks and Economic Theorizing*

The process of human learning is the foundation for making assumptions, constructing theories or models, and formulating policies. Current research, however, in parts of philosophy, psychology, linguistics, and even art history—all converge to suggest that the traditional Western either/or conceptual framework is somehow askew and no longer functions effectively [6]. In regard to economics, the preceding analysis indicates that the way in which one specifies the economic-culture relationships is consistent with one's particular conceptual approach. The question arises whether this general analysis has explanatory power when applied to the development of economic theory itself. Essentially I contend that, regardless of one's intentions, either/or thinking limits the scope of economic theorizing while the di-unital habit of mind, one of many possible alternate frameworks, enlarges it.

Consider for the moment a person who has an either/or perspective and whose particular culture predisposes him to advocate the value of people as individuals.⁷ This person cannot simul-

taneously acknowledge in a positive fashion those cultures that engender opposite values such as the notion of people as a group, people as an entity unto themselves. Recognizing only his culture, for instance the white American one, he logically considers it universal or perhaps a minor variation of some "universal" Western culture. To achieve his stage of development, it follows in his eyes that other groups such as Afro-Americans must remold themselves culturally in his universal image. Accordingly, he chooses assumptions to underlie his economic theories that are consistent with this cultural perspective. He might reason in the following manner: since each individual knows his own interest best and people as a group are nothing more than a collection of these individuals, it is clearly wisest to allow individuals to maximize their respective self-interests. In this manner, the pursuit of self-interest in the form of an independent utility function quite logically becomes a cornerstone upon which to formulate a valuable and powerful body of individual-oriented economic theory.⁸

Unfortunately the either/or perspective limits perception of and subsequent theorizing about economic behavior that reflects group-oriented cultural traits. Assume that an Afro-American, for example, obtains urban employment; thereby causing his extended family (relatives, etc.) to migrate and to share his new resources. According to the perception of the black person, satisfaction is maximized, a large measure of which includes the satisfaction of a number of people outside the immediate family (interdependent utility function). Since the either/or person cannot acknowledge in a positive way the cultural base of this behavior, he does not use the experience to formulate economic theory consistent with interdependence and group identity. Instead his conceptual framework commits him to apply individualistic economic theory as if it were universally valid. Yet he cannot do this because the marginal conditions for optimal resource allocation no longer

his needs. Accordingly, a group becomes only a collection of different individuals.

⁸ Independent utility function means that in striving for self-satisfaction, I do not consider whether other individuals are better or worse off as a result of my economic behavior. Although interdependent utility is present in this type of culture (for example, parents in a nuclear family receive satisfaction from feeding their children), independent utility functions, nevertheless, correctly characterize this culture from an aggregate viewpoint. This assumption of selfishness underlies much of standard and neoclassical economic theory of the firm and the consumer. In welfare economics it is the basis of the concept of Pareto optimality.

⁷ For example, in white American culture the corporation is legally an "individual"; the Bill of Rights is for individuals. The nuclear family structure and the institution of private property prevail. Individual effort determines one's position: Truman from the haberdashery, Eisenhower from the farm, Kennedy from Irish immigrants, Johnson from the backwoods of Texas. Each individual is defined more by the way in which he differs from other persons than by the way in which he is similar to them. One is paid according to his productivity, not

hold. The Afro-American's wages, the social marginal cost, do not adequately reflect his private marginal cost, since part of his earnings are analogous to rent paid to nonproductive factors; i.e., members of his extended family. Similarly, price, the social marginal utility of the goods purchased by the Afro-American, no longer indicates his own private marginal utility, since the latter includes the private marginal utilities of a number of other consumers. In other words, this suggests that while the desire to maximize utility functions is the province of universal economics, the types of functions are the concern of nonuniversal economics.

Alternatively, the di-unital framework suggests a different and less limiting approach to economic theorizing. In order to see this, let us consider a person with economic training—for example, an Afro-American like me—who recognizes the co-presence of opposites. In this capacity DuBois poses for me the following problem: "There faces the American Negro therefore an intricate and subtle problem of combining into one object two difficult sets of facts: his present racial segregation which despite anything he can do will persist for many decades; and his attempt by carefully planned and intelligent action to fit himself into the new economic organization which the world faces" [5]. After a life-time of studying American race relations, DuBois recommended that black people develop their own economic institutions to provide goods and services in their already segregated or separate communities. Some black people currently advocate a similar solution. Others see the solution in the formation of a black state, the latter being a more complete and formal geographical separation of the American races.

As an Afro-American I cannot dismiss these solutions out of hand; as a di-unital Afro-American with economic training, I respond to them in the following way. According to a di-unital perspective, the nature of the relationship between races is more important than whether the races are geographically separated. While a black nation may structure its institutions to reflect mainly black ethnic culture, at the same time these institutions must recognize the cultural existence of white Americans. In fact, a black nation of this type may actively seek to have whites as part of its citizenry, since black cultural uniqueness is also informed by other cultures, an important one of which is white American European. In a di-unital-oriented nation, a white person can also develop his white and nonwhite cultural components through interaction with its black citizenry. Consequently, this habit of mind allows me to see

that black American communities may be validly separate and at the same time truly connected to white America and the rest of the world.

In dealing with this solution of black economic separation, I also recognize that economics is not culture free. As a result, two points occur to me. First, the cultural rationale proposed by some black nationalists for a separate economy lies in the realm of the noneconomic welfare of black people. Given that Afro-Americans share a common heritage and that their heritage, their way of life, and ethnic identity are unique, these nationalists argue that black people should control their own social, economic, and political institutions. Black people can best maximize the destiny of black people. Second, in the area of economics, the traditional economist's concept of economic welfare is formulated in terms of individual consumption of goods and services. The reason suggested by Boulding [2, p. 6] for this individualistic focus is that economists have concentrated on exchange as the object of study, and exchange frequently occurs under conditions of at least relative indifference or selfishness. In this context, economists seek to maximize trade or real product defined as the total production of privately appropriate goods and services.

At this point, there arises the question of how to evaluate the feasibility of a separate economy for black Americans. As a di-unital economist, this proposal must be viewed in terms of maximization of a mixture of economic and noneconomic objectives. Evaluations in terms of maximization of only real product are limited in scope according to my perspective.⁹ The di-unital answer involves also attaching utility or value to the nonpecuniary objectives (defined by collective black preference functions) of a separate black economy of a nonautarkic character. Therefore, I am directed to ascertain and to evaluate the loss in economic welfare that can potentially be yielded by a separate black economy.¹⁰ Further, I must specify this relationship between the economic and

* These evaluations generally hold that the separation of races is uneconomic, since trade is diminished between communities. One reason offered is that to engage in a vast import substitution program means developing black economic institutions to replace the present white ones. In turn, the former will provide goods and services to blacks at higher costs and less efficiently than justified by considerations of productivity. A suboptimal total output results.

¹⁰ Consider, for example, the problem of allocating a given sum of money to one of two firms to build housing in a black community. One firm is low-cost, white-owned, located outside the community and can supply 23,000 housing units; the other company is

noneconomic objectives. After defining one such model, I as a di-unital black economist can then evaluate the proposal for a separate black economy; thereby producing a solution to the DuBois problem.

IV. Concluding Comments

This paper does not include an attempt to develop a body of theory that may be termed "black economics." Instead, I have set forth a conceptual framework for the formulation of economic theory that permits one to be stimulated by and to be consistent with the cultural distinctness of Afro-Americans. In turn, this approach allows for the possibility of black economics. Evaluating this possibility or developing a body of black economic theory is not necessarily the task of black economists alone. For example, white American economists who have a real understanding of black culture can also make valuable contributions. This understanding, however, must be based on extensive exposure to and training in black cultural patterns similar to the way in which some black people have learned white

American European patterns. In other words, for white people who think in di-unital terms, our blackness is part of their "not being white" component.

I also recognize that cultures similar but not identical to black culture may also provide the same stimuli to the development of economic theory. It is not my prime concern to generate a problem of cultural identification in the economics discipline. An understanding of the proposition, economics is economics, in a di-unital sense recognizes sufficiently the role of the nonuniversal economics, one of which may be a body of black economic theory. I am aware that the di-unital approach raises fundamental epistemological problems that I don't pretend to have answered here. Furthermore, I have only begun to work out the economic implications of this approach. A more complete undertaking must wait until another occasion.

REFERENCES

1. S. BARATZ AND J. BARATZ, "Early Childhood Intervention: The Social Science Base of Institutional Racism" (forthcoming in the Winter, 1970, *Harvard Educa. Rev.*).
2. K. E. BOULDING, "Economics as a Moral Science," *A.E.R.*, Mar., 1969.
3. V. J. DIXON, "The Black Students and the Brother in the Street," *Negro Digest*, Nov., 1968, pp. 28-35.
4. V. J. DIXON AND M. R. LEWIS, "Black Consciousness, Societal Values and Educational Institutions: A Statement of Linkages," *J. of Conflict Resolution*, Sept., 1968, pp. 402-07.
5. W. E. B. DuBois, *Dusk of Dawn* (Harcourt, Brace and Co., Inc., 1940), p. 199.
6. T. S. KUHN, *The Structure of Scientific Revolutions* (Univ. of Chicago Press, 1962), pp. 120, 125.
7. B. RUSSELL, *The History of Western Philosophy* (Simon and Schuster, 1945), pp. 731 ff.

high-cost, black-owned, located within the community and can supply 21,000 housing units. To maximize satisfaction from the consumption of housing alone, the white-owned firm should receive the allocation. It uses the given amount of money more efficiently; i.e., it supplies more housing than the black-owned firm. This approach fails to consider explicitly that the collective black preference for having its own productive facilities may be a consumption good which yields satisfaction independent of the consumption of housing itself. This latter gain in consumption may be sufficient to compensate for the loss in housing units produced by the allocation of funds to the black-owned firm, and therefore may justify its receiving the allocation.

THE INTERNATIONAL FIRM AND EFFICIENT ECONOMIC ALLOCATION

INTERNATIONAL TRADE IN INPUTS AND OUTPUTS

By ROBERT E. BALDWIN
University of Wisconsin

I

In defending the twin assumptions that commodities can move freely among countries but primary factors are completely immobile internationally, trade theorists generally point out that without the factor-immobility assumption the distinction between international trade theory and domestic production and exchange theory disappears.¹ However, since the regional pattern of trade as well as the geographic distribution of productive factors becomes indeterminate when it is assumed that both goods and factors are perfectly mobile within a country, domestic production and exchange theorists usually assume all economic activity takes place at one point in space. Consequently, factor movements and their interrelationships with commodity flows have not been analyzed within the mainstreams of either international trade theory or domestic production and exchange theory. Instead, the subject has become a subsidiary topic of economic theory—studied mainly by location theorists, by economic historians, and, more recently, by economists interested in development theory.

The leading trade economist who has tried to change the typical practice of separating the treatment of commodity and factor flows is, of course, Bertil Ohlin.² As he states in the Preface, a major purpose of his treatise is: "To analyze the domestic and international movements of factors of production, and particularly their relation to commodity movements."³ Although Ohlin's work is rich in insights on this subject, the general impact of his work has, ironically, been to reinforce the traditional approach of trade writers. For although Ohlin stressed that labor and capital are neither completely mobile or immobile internationally, he in effect assumed in his simplified trade model that knowledge was completely mobile and, therefore, that production functions were everywhere the same.⁴ It then re-

mained for Samuelson to show that, by adding a few seemingly reasonable assumptions, factor prices become equalized through trade.⁵ Despite Samuelson's warning that the actual disparity in factor prices among countries meant that these assumptions were not so innocuous after all, the factor-price equalization model has tended to become the cornerstone of international trade theory. And, since the same world production possibilities are attainable in this model with commodity trade alone as with commodity plus factor trade, the tradition of ignoring factor movements has been further justified.

Recent events, especially in connection with the operations of international firms, have, however, made it increasingly inappropriate to ignore the interrelations between output and input flows. Trade economists could in the past partly justify their position on the grounds that different decision-making units were usually involved in commodity and factor flows and that in the nineteenth century a large share of factor flows were directed at the production of noninternationally traded services, e.g., canal and railway services, or of commodities effectively unavailable in the developed countries, e.g., tropical products and certain minerals. But, today we frequently observe the phenomenon of an international firm weighing the alternatives of producing a particular commodity in one country and then shipping it to the market of another country or transferring technology and productive factors to this latter country and manufacturing the product there. The possibility of various patterns of trade in intermediate inputs makes the set of feasible alternatives facing the international firm even more complex.

In order to understand better the nature of current international commodity and factor flows and to be able to deal more adequately with the policy issues they raise, we should return to Ohlin's broad vision of studying these flows simultaneously. It is also important that we consider the institutional form that these flows take. Fortu-

¹ See, for example, G. Haberler, *The Theory of International Trade* (London: William Hodge, 1936), pp. 4-5.

² Bertil Ohlin, *Interregional and International Trade* (Harvard Univ. Press, 1952).

³ *Op. cit.*, p. viii.

⁴ *Op. cit.*, p. 557.

⁵ Paul A. Samuelson, "International Trade and Equalization of Factor Prices," *Econ. J.*, June, 1948.

nately, more and more work along these lines has been taking place in trade theory, but we are still severely constrained by the traditional approach.⁶ In this paper I should like to outline a simple, Ohlin-type framework for analyzing this topic and then, drawing upon existing knowledge, discuss very briefly some of the main factors that influence the pattern and magnitude of international input and output flows as well as some of the interaction effects of these flows.

II

A useful starting point for a model of trade in outputs and inputs is the simple notion that a country's demand for a particular commodity can usually be met in any one of three ways: by producing the commodity entirely within the country from existing primary inputs; by importing the final product from other countries; or by importing some of the primary and intermediate inputs needed to produce the item. The specific manner chosen to meet the demand depends not only on the relative costs among countries of producing commodities and hiring factors but upon the costs of transferring outputs and inputs (including knowledge) from one country to another. Trade theorists have traditionally stressed the transportation component of transfer costs. However, another very important part of transfer costs that economists recently have begun to emphasize are information costs.⁷ It is costly for commodity-producers, owners of productive means and final consumers of products and factors to obtain and collate information about exchange opportunities. Moreover, like transportation costs, information costs vary among products and factors and follow the usual laws of production costs.

Since technological knowledge is not usually freely transferable among countries, production functions are not assumed to be necessarily identical among countries in the model being suggested here. It is also inappropriate in analyzing input and output flows to assume that all production functions exhibit constant returns to scale. However, it appears that most of the other assumptions of the standard Heckscher-Ohlin-Samuelson trade model—e.g., diminishing marginal produc-

tivity, uniform factor-intensity ratios, qualitatively identical primary inputs, a greater number of outputs than inputs (but not that some quantity of all commodities is produced in each country)—can be usefully accepted in analyzing these flows.⁸

International trade in commodities frees a country from the constraints imposed by its own factor and knowledge endowment and permits it to take advantage of other factor endowments that are comparatively better suited to produce particular commodities. Commodity plus factor trade goes a step further and removes the restriction of a fixed and limited set of relative factor endowments in the world economy. If there are no transfer costs for inputs or outputs and if there are identical production functions that are linear and homogeneous, world income can be maximized either by commodity trade alone or by factor movements alone (or by some combination of the two). However, if transfer costs for both products and primary factors exist and production functions are not identical among countries, both a rearrangement of the stock of factors and commodity trade based on this new factor allocation is generally needed to achieve a long-run, maximization of world income.⁹

III

Two general sets of conditions determine the nature of output and input flows among countries in the type of model outlined here. One concerns differences among countries in factor endowments, technology, and tastes, and the other relates to the various factors affecting the costs of transferring outputs and inputs internationally. Little will be said about the first set of conditions except to note that both the output and input flows induced by intercountry differences in relative factor supplies and technology can be usefully analyzed under the familiar twofold classification of "push" and "pull" forces. Depressed domestic product-markets, unused productive resources within domestic firms, or direct invest-

⁶ The usual assumption concerning perfect factor mobility and commodity mobility within each country is also accepted.

⁷ In addition to the usual domestic and balance-of-payments conditions, international equilibrium under perfectly competitive market conditions requires that the difference in the price or capitalized value of any output or input between two countries is equal to or less than the marginal costs of transferring the output or input between the two countries. With fixed factor endowment and knowledge conditions, the long-run equilibrium position will be one in which there are no factor flows among countries.

⁸ A good example of this shift in emphasis recently is the change given by Kemp to the title (and content) of his monograph on trade theory from *The Pure Theory of International Trade* to *The Pure Theory of International Trade and Investment* (Prentice-Hall, 1969).

⁹ An important recent article on the subject is Armen A. Alchian, "Information Costs, Pricing, and Resource Unemployment," *Western Econ. J.*, June, 1969.

ment by competing firms in foreign markets are examples of "push" forces. "Pull" forces, on the other hand, include such well-known conditions as an increase in wage rates and in employment opportunities abroad or the lowering of foreign tariff barriers. Needless to say, a wide variety of input and output flows can be generated by these forces. In particular, all physically movable factors may flow in one direction when trade is opened up between two countries or there may be an exchange of factors as well as of commodities between the two economies.

It is the second set of conditions determining the nature of output and input flows—namely, the various factors affecting the relative costs of transferring commodities and factors between countries—that traditional trade theory tends to ignore and that, therefore, this paper focuses upon. Output and input flows between two countries are influenced, not only by the absolute size of the ratio of transfer costs to the value of the outputs and inputs being sent abroad, but also by the relationship between the transfer costs on commodities and the transfer costs on the inputs used to produce the commodities. For example, one familiar set of considerations influencing both the transportation and information costs connected with international flows of outputs and inputs is the physical characteristics of different outputs and inputs. As location theorists have long pointed out,¹⁰ commodities that are comparatively heavy or bulky in relation to their value tend not to move in international trade except where the costs of transporting the primary inputs needed to produce the commodities, e.g., natural resources, are even relatively higher. Similarly, the movement of low-wage labor, which involves relatively high transfer costs, has frequently occurred under the circumstances where the production process also involved the use of a rich but immobile supply of natural resources. Commodities that can be easily graded or standardized have a mobility advantage over heterogeneous products with comparable production costs because of their relatively lower information costs. In the capital goods field, for example, the specialized nature of many product requirements, coupled with the high costs of maintaining comprehensive inventories of a particular product line in order to reduce delivery times, tends to fa-

vor the production of such products near the market where they are sold.¹¹ The communications advantage provided by physical proximity between capital goods producers and consumers seems to play an especially important locational role in technologically dynamic product lines.¹²

Since there are significant economies of scale involved in transferring factors and products, the size of the foreign market for particular inputs and outputs is another important factor affecting relative transfer costs. When, for example, an employment market is small, individual employers as well as potential migrants abroad are faced with very high costs of providing and obtaining information about job opportunities. However, unit search costs decline as the market expands and specialized recruiting organizations and methods become economically feasible. The development of capital markets is another institutional arrangement that reduces transfer costs; yet is only economical if the market opportunities for capital are relatively large.¹³

Limited product markets also impede the international flow of commodities. However, for many goods the high costs of shipping the product to a small market is less of a trade impediment than the high costs of transferring inputs and undertaking production in a small market. Consequently, a typical sequence of trade flows has been first commodities and then productive factors as the market expands and as the commodity flow provides low-cost knowledge to both domestic and foreign entrepreneurs about investment opportunities in the particular product line.

Producers and productive agents benefit (or sometimes are harmed) from a variety of externalities as foreign markets for a country's inputs and outputs expand. The spillover effects associated with labor migration are a case in point. As the number of migrants to a foreign country increases, the knowledge about additional employment opportunities obtained by these migrants during the normal course of their work and their social contact with other migrants is sent back to friends and relatives in the "home" country at little or no cost.

¹⁰ Jack N. Behrman, *Some Patterns in the Rise of the Multinational Enterprise* (Graduate School of Bus. Admin., Univ. of North Carolina, 1969), p. 3.

¹¹ Nathan Rosenberg, "Economic Development and the Transfer of Technology: Some Historical Perspectives," *Technology and Culture*, July, 1970.

¹² These are illustrations of the general point made by Allyn Young, "Increasing Returns and Economic Progress," *Econ. J.*, Dec., 1928, and elaborated upon by G. Stigler, "The Division of Labor Is Limited by the Extent of the Market," *J.P.E.*, June, 1951.

¹³ As Ohlin appreciated (*op. cit.*, Chap. XII) together with modern location theorists such as W. Isard (*Location and Space-Economy*, Wiley, 1956, especially Chap. 9) trade and location theory tend to merge under the trade framework being suggested in this paper.

This significantly reduces the search costs of this latter group and stimulates their movement abroad. Similarly, information about investment opportunities in a particular industry abroad is often obtained at relatively little cost by firms in the same industry or with close backward or forward input-output linkages to the foreign industry. Consequently, foreign investment by these firms is facilitated. Purchases by investors and migrants of the home country products with which they are familiar also have the effect of spreading information about home products to other residents abroad and thus of increasing the sales of these products. Still another spillover benefit to individual producers and factors is the free information about employment and investment opportunities that governments and news media provide.

The degree of risk and uncertainty present in foreign markets is another factor affecting international input and output flows. A high degree of risk and uncertainty acts to discourage all types of input and output flows, but this condition particularly impedes capital movements due to the indivisibility and comparative immobility of most specific capital goods in which foreign capital funds usually become embodied. Consequently, commodity flows and, to some extent, movements of highly skilled, relatively mobile labor are favored in risky markets as compared to capital flows.

In addition to the size and riskiness of product and factor markets abroad, a significant determinant of the nature of input and output flows to these markets is the optimum size of the decision-making units that send and receive these flows. In industries where the basic production unit is sufficiently large so that the problem of management indivisibilities can be overcome and thus efficient international firms established, a wide range of opportunities for reducing costs by means outside of the open market mechanism become available in the industry, especially in connection with what Edith Penrose terms "economies of growth."¹⁴ For example, because of lower costs of obtaining information about investment prospects and of raising and transferring funds abroad, it tends to be cheaper for large-scale, growing productive units in one country to obtain funds from commonly-owned and managed productive units in the same industry in another country than to obtain these funds from the capital market of this latter country. However, probably the most significant growth benefit associated with the international firm relates to the transfer of knowledge among

countries.¹⁵ Transferring technological knowledge through the market mechanism, e.g., by patent-leasing arrangements, is often not very efficient, especially if the levels of knowledge and skills in the recipient area are low. In such cases, the very lack of knowledge on the part of the recipients prevents them from appreciating the full value of the knowledge available from more technically advanced firms abroad. More important, it is usually necessary to send a team of technical experts to the recipient firm temporarily in order to transfer the technology successfully. Thus, the more complex the technology to be transferred in relation to the level of technical skills in the recipient firm and the more rapid the flow of new products and productive methods from the more advanced country, the more efficient it becomes for firms in the technically advanced country to transfer their knowledge by establishing foreign subsidiaries. The fact that much of the knowledge is specific to particular individuals, who acquire their value only as members of an integrated technical team, also reinforces the advantages of transferring knowledge through the international-firm mechanism.¹⁶ Because of the specific knowledge about dealing with customers, labor organizations, and the government that existing firms in the technically less advanced countries possess, firms in the more advanced countries also usually find it cheaper to merge with or acquire the former firms rather than start completely new operations.

When the international firm becomes economically viable in a particular industry, not only is it possible to transfer knowledge, capital, and technical and managerial labor across borders more efficiently, but these transfers tend to be economically feasible with smaller product markets than is the case when the optimum size of productive units is small. Moreover, because of the pecuniary and technological externalities that exist among intermediate sectors in a vertical product line, it may not be economically profitable from a private viewpoint to add a new product to a country's production list unless the international-firm mechanism is utilized.

A final factor that should be mentioned in discussing the costs of transferring outputs and inputs among countries is the nature of technological developments in transportation and communication. Recent advances, especially the introduc-

¹⁴ N. Rosenberg, *op. cit.*, p. 11, makes this point. Raymond Vernon analyzes the economies of scale involved in the production of technological knowledge in an unpublished paper, "Organization as Scale Factor in the Growth of Firms."

¹⁵ These points are discussed much more comprehensively in Jack Baranson's paper.

¹⁶ Edith T. Penrose, *The Theory of the Growth of the Firm* (John Wiley, 1959), pp. 99-103.

tion of the jet aircraft, have favored the increased flow of high-valued (per unit of weight) commodities and factors. In particular, it has significantly reduced the opportunity costs of sending highly paid technicians and managers abroad for short periods and thereby enabled the international firm to become economically feasible with smaller productive units. Export-oriented foreign investment in small-scale manufacturing lines that involve a high-value product and a high labor-capital ratio have been greatly stimulated in recent years by such developments in transportation and communications. The advances in transportation and communications technology have also tended to speed up technological change itself by increasing the amount of personal contact and confrontation among scientists, engineers, and other professionals.

IV

The nature of international factor flows is very different today than from the nineteenth century. In the last century capital—and to a considerable extent labor—moved into activities that directly or indirectly increased exports of natural-resource products. Exports of the main capital-supplying countries were manufactured goods, and these tended to rise as a result of the factor flows. In recent years advances in production and managerial technology together with technological improvements in transportation and communications have greatly increased the potential mobility of capital, technological knowledge, and professional and technical labor. The mobility of unskilled and semiskilled labor, on the other hand, has been sharply reduced by artificial barriers. Although considerable foreign investment by the developed countries is still directed at increasing their supplies of natural-resource products, a growing share of factor-export activities by developed countries, particularly the U.S., is based on their specialization upon the production of technological knowledge in the manufacturing field. On an increasingly large scale, technology, highly skilled labor, and capital flows abroad from these countries via the international firm and results in the establishment of manufacturing units that com-

pete directly with domestic export- and import-competing production. These developments are mainly based on real factors but they have also been stimulated by immigration restrictions on unskilled labor imposed by the developed countries and by commodity barriers maintained by developed and developing countries alike.

The effects of these manufacturing activities in diffusing information about other products and reducing transfer costs as well as in bringing about favorable price and income repercussions abroad may be sufficient to maintain aggregate exports of manufactures from a country like the U.S., but the composition of the country's trade is likely to change drastically. In particular, exports of less technologically dynamic product lines are likely to decline sharply (and imports of these items rise) and the labor force producing these commodities put under severe income pressure.

The economic viability of the international firm is based to a large extent on economies of growth, especially if there is a large initial technological and human-capital disparity between countries. As the gap in technological and skill levels narrows between countries in particular industries, one would expect that international firms in these industries would tend to become looser multinational organizations or, in some cases, to be broken up into national firms. However, the unique market power possessed by these firms may be used to resist such socially-beneficial changes.¹⁷ We have developed a nonmarket mechanism for facilitating economic growth in the world economy, but our internal and international institutions for coping with the economic shocks resulting from this growth and with the possible socially undesirable economic and political effects of the internationalization of production are in need of considerable improvement. If international trade theorists are to contribute towards this improvement, they must cast their analysis in a framework that includes trade in both outputs and inputs among countries.

¹⁷ Stephen Hymer, "Direct Foreign Investment and the National Economic Interest" (Economic Growth Center, Reprint Paper No. 108, Yale Univ.).

TECHNOLOGY TRANSFER THROUGH THE INTERNATIONAL FIRM*

By JACK BARANSON

International Bank for Reconstruction and Development

Technical knowledge has become an important item of international trade, and, as with all scarce resources, there are deficit and surplus nations in the world economy. In 1964, international payments for technology reached the one billion dollar mark, with the United States and Western Europe as principal suppliers. These payments were for royalties and licensing fees on patents and know-how covering products designs and production techniques.¹ Successful transplants of technology depend as much upon willingness and ability to transfer technical knowledge and skills as upon the absorptive capabilities of recipients. In the case of less sophisticated partners, successful transfers generally require a sustained relationship between donors and recipients. This includes the transplanting of managerial systems governing quality control and materials standards, production scheduling, materials handling, procurement, and labor productivity.

In the past two decades, there has been a tendency toward manufacturing overseas as a precondition to serving national markets. This trend has given rise to multinational production and interchange, which intensifies the need for effectiveness in technological transfer as a contributor to efficiencies in resource utilization. Transfer across national boundaries, as distinct from enterprise-to-enterprise transfer within national boundaries, is

perceived and valued differently by transmitters of technology because of differences in legal systems, conflict of political interests, and the added uncertainties of changes in exchange rates. Modes of transfer range from direct investment, with full ownership and control, to joint ventures and licensing of patents and know-how on a contract-fee basis. This paper focuses upon the factors influencing transfer logistics and the conflicts of interests in licensing versus investment decisions.

Factors Affecting Transfer Logistics

The relative cost and feasibility of technology transfer depends upon four sets of interrelated factors: (1) the complexity of the product and production techniques being transferred, (2) the transfer environment in the donor and recipient countries, (3) the absorptive capabilities of the recipient firm, and (4) the transfer capability and profit-maximizing strategy of the donor firm. Decisions by international firms favoring either licensing or direct investment are strongly influenced by a combination of these factors.

1. *Nature of the Transferred Technology.* Various characteristics of product design and the relative sophistication of production techniques have a bearing upon the investment-licensing decision. Donor firms are less interested in licensing the product fields where substantial R and D funds have been invested or where they hold a technological lead and have the resources to undertake direct investment. However, the shorter the product cycle, as in microelectronics and pharmaceuticals, the more willing a firm may be to license in residual markets that they are not in a position to exploit through direct investment. In new or highly sophisticated technologies, which may involve considerable teething trouble, the tendency is toward direct investment and a sustained relationship with the overseas affiliate. The same applies to certain products and processes, such as shuttleless looms, tunnel brick kilns, and continuous casting techniques—all of which require refined adjustments to local materials and customer preferences.²

2. *Transfer Environments in Donor and Re-*

* Obstacles to diffusion in these fields are described and analyzed in G. F. Ray, "The Diffusion of New Technology," *Nat. Inst. Econ. Rev.*, May, 1969, pp. 40-83.

* This paper is based in part upon correspondence and interviews with some thirty corporate executives of international firms with overseas manufacturing affiliates. I am especially grateful for comments and suggestions received from John Tilton, G. F. Ray, Richard D. Robinson, Walter Chudson, and Judd Polk.

¹ The United States received 57 percent and Western Europe 41 percent of the world's \$1 billion in technological payments in 1964. Japan paid 18 percent, as compared to 11 to 14 percent each by the United States, the United Kingdom, Germany, and France, and all developing countries paid a total of 8 percent. (It should be noted that these figures may represent only a fraction of the value of acquired technology, which may be paid for in the form of cross-licensing, remitted profits on equity holdings in foreign affiliates, or earnings from sales of equipment and parts to licensees.) See United Nations Conference on Trade and Development, *Trends and Problems in World Trade and Development*, study prepared by C. H. G. Oldham, C. Freeman, and E. Turkcan, for the Second Session UNCTAD, New Delhi, India, Feb. 1, 1968, TD/28/Supp. 1 (mimeo., Nov. 10, 1967), p. 13.

Recipient Countries. Licensing versus investment decisions by international firms are influenced by a variety of legal constraints in both the donor and recipient countries, affecting the transfer and ownership of investment resources. The competitive position of a firm at home and abroad (which in turn may depend upon uniqueness of the firm's product, its corporate skills, and market accessibility) also has a strong bearing on the firm's willingness to license versus an insistence upon equity shares. Laws controlling foreign investment, anti-trust provisions in the U.S. particularly, and tax structures at home and abroad also affect choices between licensing and investment.

International firms experience particular difficulties in developing economies, which stem from the early stages of industrial development and the limited size of internal markets, coupled with economic policies of forced industrialization and a general tendency toward economic nationalism. Developing economies generally lack the full range of skills needed to run a successful manufacturing operation, but insistence on national ownership and staffing often inhibits even the interim use of foreign technicians in vital areas of plant operations. Policies of protection and high domestic content also create serious problems in local procurement. More recently, governments in developing countries have been urging local manufacturers to produce for export markets, which places an even heavier burden on technology transfer agents in terms of quality and cost competitiveness. Exchange controls associated with industrialization policies and balance-of-payment difficulties add to the problem of running plants in developing economies and remitting equity earnings or royalties and licensing fees.³

3. *Recipient Firm's Characteristics.* In considering whether to enter into a license arrangement as an alternative to direct investment, the technical absorptive capability of the recipient firm and its potential in competitive markets are major factors.⁴ For example, licensing arrangements are of-

ten preferred with industrially-advanced partners, because of the ease of transfer coupled with the advantages of cross-licensing. Direct investment is sometimes preferred in underdeveloped countries because of the uncertainties and added difficulties in imparting technology in those areas. Donor firms are also more willing to disclose technical know-how to less sophisticated partners in developing countries than they are to industrially-advanced firms, which may eventually become serious commercial rivals in third markets. Where the donor firm has a strong technological lead and a dynamic R and D program to maintain that lead, this consideration may be less important.

4. *Donor Firm Characteristics.* The transfer capability, financial position, and corporate philosophy of the donor firm are among the decisive factors influencing licensing-investment decisions. International firms may realize returns on their technological assets in a variety of ways: dividends on equity investment, sale of components and parts, royalties, licensing fees, and technical assistance fees. Direct investment involves the commitment of financial and managerial resources to which there are limits, even for the largest corporations. For example, in the automotive field, international firms have had to develop low-cost systems to package and ship components and provide cadres of managers and technicians to staff overseas assembly and manufacturing operations in newly industrializing countries. Even straight licensing arrangements often require commitment of managerial resources and the development of new corporate capabilities, if product standards are to be maintained internationally. Corporate views on how to maximize long-run profits (costs of developing transfer capabilities as against benefits from licensing versus direct investment) have a strong bearing upon the chosen transfer mode. Certain firms view technology as an integral part of their marketing-manufacturing package and insist on full ownership and control to maximize international profit.⁵

³ For an analysis of encountered difficulties, see author's *Manufacturing Problems in India: The Cummins Diesel Engine Experience* (Syracuse Univ. Press, 1967); and author's *Automotive Industries in Developing Economies*, World Bank Staff Occasional Paper No. 8 (Johns Hopkins Press, 1969).

⁴ For an analysis of differences in absorptive capabilities among Japanese component and subassembly manufacturers in the jet aircraft field, see George Hall and Robert Johnson, "Transfer of U.S. Aerospace Technology to Japan," paper delivered at the Conference of Technology and Competition in International Trade, sponsored by the Universities-National Bureau Committee on Economic Research

(New York, 1968). For distinctions made in foreign licensing arrangements with firms of differing competitive potential in the diesel engine manufacturing field, see author's *Manufacturing Problems in India*, *op. cit.*, pp. 24-28.

⁵ On the trend toward global decisions on the location of production, product mix, supply procurement, interaffiliate pricing, and the coordination of marketing, research, and financial programs, see Jack Behrman, *Some Patterns in the Rise of the Multinational Enterprise* (Univ. of North Carolina, 1969), pp. 62-94. For an analysis of why international firms seek to control their international operations, see

Conflicting Interests over Licensing versus Investment Choices

International firms prefer direct investment to licensing where: (a) the financial and human resources are available, (b) control over present and future market developments is desirable, particularly with products and techniques having a longer life cycle, (c) the firm fears licensing will result in the give-away of valuable know-how or will threaten its position in established markets, (d) the transfer involves a broad line of products or is an integrated part of marketing and financial management, (e) the technology is highly complex or the foreign affiliate lacks industrial sophistication and the transfer requires a prolonged and sustained relationship to effect the transfer, or (f) there is a concern over protecting the product standards or trade name. Also, for U.S. firms, transfer to controlled subsidiaries avoids most antitrust aspects entailed in licensing to third parties.

International corporations favor licensing over investment where: (a) the market is too small to warrant investment or the product cycle or proprietary position is ephemeral; (b) the firm has a marketable technology, but lacks the resources or experience for direct involvement; (c) direct investment is precluded by legal constraints or seems to involve high risk and uncertainty of a political or economic nature; (d) reciprocal benefits are obtainable through cross-licensing; or (e) patent litigation or competitive technological development may be avoided.⁶

The major conflicts over technology transfer between donor firms and recipient firms (and their governments) concern pricing, ownership,

and long-term technological development. The pricing of technology and the mode of transfer (licensing versus investment) are essentially determined by the degree of competition among donor firms, the relative bargaining power of the contracting parties, and the regulations imposed by recipient governments.⁷ In Japan, for example, there is a purposeful effort to put ownership and control of acquired technology in Japanese hands and to nurture indigenous technological absorptive and innovative capabilities. For these reasons, licensing is the preferred mode of technological transfer on imported technology. The Japanese government also reviews licensing and investment agreements to assure there are no unreasonable constraints on its future use for trade in world markets.⁸

Many developing countries now share a similar concern about developing indigenous technological capabilities and prefer licensing or majority share indigenous ownership to foreign control. They are also resentful of the price charged for acquired technology.⁹ It has also been argued that foreign corporations often remit earnings from manufacturing operations based upon domestic resources, to which a thin veneer of foreign investments in technology and capital has been added. These high returns are augmented by overvalued exchange rates that are typical of developing countries.

In many instances, governments in developing

Charles P. Kindleberger, *American Business Abroad* (Yale Univ. Press, 1969), pp. 1-36. See also Sidney E. Rolfe, *The International Corporation*, report prepared for the XXII Congress of the International Chamber of Commerce, Istanbul, May 31-June 7, 1969 (I.C.C., 1969); and Sanford Rose, "The Rewarding Strategies of Multinationalism," *Fortune*, Sept. 15, 1968, pp. 96-183.

⁶The pros and cons of domestic and foreign licensing practices from the commercial and legal standpoints are analyzed in Joseph S. Cardinale, *Manual on the Foreign License and Technical Assistance Agreement* (Thomas Ashwell, 1958); K. S. Goldschmid, *International License Contracts: A Practical Guide* (Copenhagen: Ministry for Foreign Affairs, 1968); E. B. Lovell, *Domestic Licensing Practices: Experiences in Marketing Management* (National Industrial Conference Board, 1958); E. B. Lovell, *Appraising Foreign Licensing Performance* (National Industrial Conference Board, 1969), and G. M. Pollizien and G. B. Bronfen, eds., *International Licensing Agreements* (Bobbs-Merrill, 1965).

⁷Factors influencing decisions by international corporations relating to technology transfers are analyzed in G. C. Hufbauer, *Synthetic Materials and the Theory of International Trade* (London: Gerald Duckworth, 1966), pp. 89-92; Robert B. Stobaugh, Jr., "Where in the World Should We Put that Plant?" *Harvard Bus. Rev.*, Jan.-Feb., 1969, pp. 129-36; Raymond Vernon, "International Investment and International Trade in the Product Cycle," *Q. J. E.*, May, 1966, pp. 190-207; and Richard D. Robinson, "Ownership Across National Frontiers" (working paper prepared for the Sloan School of Management, Massachusetts Inst. of Tech., Feb., 1969, mimeographed).

⁸On the impact of government controls upon the acquisition of foreign technology, including the indigenous development of technology in Japan, see Terutomo Ozawa, "Imitation, Innovation, and Trade: A Study of Foreign Licensing Operations in Japan" (doctoral dissertation, Columbia Univ., 1966).

⁹On the growing concern of developing countries over the costs and constraints imposed upon transferred technology, see Miguel S. Wionczek, *Arrangements for the Transfer of Operative Technology to Developing Countries*, Preliminary Report for a Case Study of Mexico, prepared for the United Nations Economic and Social Council (E/4452/Add.3, Mar. 6, 1968).

countries have taken the view that it is unreasonable for technology donors to extract a rent over and above costs for services rendered in technology transfer. Most governments in developing countries tend to underestimate the run-in problems incurred in implanting acquired technology. These governments have argued that foreign investors should "place less reliance on the profit motive" and "should consider technical know-how as something to be shared with developing countries."¹⁰ The more adverse the environment and the more complex the technology, the more resources it takes to effect the transfer and the less willing or able a corporation may be to effect the transplant, unless adequately compensated.

Developing economies pose a special set of dilemmas for international firms. In many instances, these firms feel that it is only through the intensive and sustained relationship associated with significant ownership and control that technology can be transplanted effectively and that they can earn an adequate return. Implicit here are certain assumptions, real or imagined, that equity ownership means control of factors affecting long-term profit maximization, such as inter-affiliate pricing, production and marketing logistics, and the maintenance of product standards. They feel under normal licensing arrangements, it is much more difficult to earn what is considered an adequate return from royalties and technical assistance fees, especially if the firm is to maintain its product standards and protect its trade name. Some firms have solved this dilemma by licensing know-how under another local product trade name and in effect accepting a dual standard in the quality and performance of its products. (Caterpillar Tractor does not grant the use of its name on products built by anyone other than its own subsidiaries.) In certain processing and engineering industries, many firms have adopted a policy of only transferring less sophisticated or stable technology and equipment, and to minimize technological experimentation.

Willingness to accommodate technology purchaser's demands for licensing versus investment arrangements depends in part upon the donor firm's success in home markets and long-term growth strategy. This includes a willingness to commit financial resources and develop transfer

capabilities. Small firms rarely can afford these commitments, and large firms are sometimes unwilling to adjust corporate philosophy and practice to new patterns of world enterprise. The policies and practices of certain firms are cited at this point to demonstrate the generalizations that have been drawn thus far.

In the development and manufacture of semiconductors, there are four firms that dominate the field: Bell Laboratories (Western Electric), Fairchild Camera and Instrument, Texas Instruments, and Motorola.¹¹ Each firm has a somewhat different transfer policy reflecting difference in corporate capabilities, market position, and profit-maximizing strategies. Product designs and production techniques in the semiconductor field are intricate and complex, and keys to commercial success are: (a) manufacturing at competitive costs; (b) access to state-of-the-art technology, which may become obsolete in three to five years; and (c) a continuing R and D effort to provide a favorable cross-licensing bargaining position. Patents are often filed as a bargaining device for obtaining cross-licensing arrangements or gaining entry to a market on advantageous terms.

Bell Labs (Western Electric) has a liberal licensing policy, at home and abroad, the basic objectives of which are to avoid patent litigation and to obtain the benefits of cross-licensing. In cases involving unilateral licensing, somewhat higher royalties are charged. Antitrust considerations also demand an open-licensing policy. Texas Instruments prefers to utilize its patents and know-how overseas through wholly-owned subsidiaries. They believe that in the fast moving electronics industry, maximum profit from new technological advances can best be realized through close proprietary control, which maximizes penetration in overseas markets and provides opportunities for cross-fertilization of ideas. Licensing is also used as a residual means for income and cross-licensing opportunities. Fairchild cross-licenses new technology to other American firms, but prefers direct

¹⁰ Government of Pakistan, Department of Investment Promotion and Supplies, *Development of Private Foreign Investment in Pakistan*, document prepared for the U.N. Panel Conference on Private Foreign Investment in Developing Countries, Amsterdam, Feb. 16-20, 1969 (Karachi: Government of Pakistan, 1969).

¹¹ Bell Labs has pioneered basic technology and owns strategic patents and know-how in the semiconductor field, including the "diffusion process," which is used in manufacturing transistors and integrated circuits. Fairchild developed a special technique known as the "Planar process" to mass-produce crystalline devices at low unit costs. Crystals of varying quality are grown in batches over a six-week period—the trick is to get a high yield of high-quality crystals. Texas Instruments also holds various patents in the production of semiconductor crystals along with basic patents for the invention of integrated circuits. (See forthcoming study by John Tilton, Brookings Institution, "Diffusion of Semiconductor Technology.")

investment through wholly-owned subsidiaries outside the United States. They adapt product designs and production techniques abroad to assure maximum return from overseas markets. Motorola transmits manufacturing know-how on various devices and circuits to wholly-owned foreign subsidiaries in three stages, beginning with performance testing and marketing of assembled products, then assembly of products, and finally production of basic materials used in semiconductors.

The development and manufacture of computers is another field subject to rapid technological change. At International Business Machines, computer technology has evolved over the past ten years from vacuum tubes to transistors and then integrated circuits. As IBM's technological base has grown larger and more complex (they now have seventeen plants in fourteen countries), direct investment has increased proportionately in order to recoup their substantial R and D investments and maintain targeted growth rates. Corporate income is derived from a package of services, which includes building and designing computers, equipment manufacture, and programming and application of computer technology. IBM insists on 100-percent ownership, so as not to dilute corporate decisions on pricing, profit remittances, and international allocation of resources. After prolonged negotiations, the Japanese government finally agreed to allow 100-percent ownership, mainly to acquire computer-application techniques, in which the other three or four computer firms now in the Japanese market are weak.

IBM also has a wholly-owned subsidiary in India, which produces a product line for both the domestic and export markets. About 20 percent of a medium-size computer is manufactured locally (assembly, grey iron casting, some machining of parts, and wiring of electric circuits). The Indian plant also manufactures for export an inexpensive key punch machine with nearly 70 percent local content. The Indian government has been urging IBM to extend its manufacturing operations in the small-size computer field for export. It is a fundamental principle at IBM to maintain technological parity (which includes product designs and quality standards) in all overseas manufacturing operations. The Indian government has also been unsuccessful in persuading IBM to reduce its 100-percent ownership.

A large American subsidiary in the petrochemical field establishes overseas manufacturing affiliates where their technology has a competitive advantage over local products and/or processing techniques. They prefer direct investment and full subsidiary control in order to fully exploit

their product line. They are especially concerned over proprietary rights in new product areas involving secret know-how that has limited patent protection. The firm is willing to join with local partners that afford a unique access to the local market or provide needed local capital. They only license technology which does not compete with their own commercial plans to exploit a market through equity investment or which has no near-term application. They cross-license in industrially-advanced markets to avoid patent litigation or duplicative R and D expenditure.

Ciba, a Swiss manufacturer of chemicals, plastics, and pharmaceuticals, also prefers direct investment and complete control over its associated companies. They are reluctant to license technical processes, but are prepared to establish joint ventures where necessary. This firm is more likely to license technology resulting from screening and testing than products and techniques derived from basic research.

Pilkington Brothers, a British firm that developed a new technique for manufacturing flat glass, has a technology transfer policy which is largely a function of world market structure and corporate philosophy. The "float glass process" which they developed is an advanced technique for manufacturing high-quality flat glass of varying thickness on a large scale at costs considerably lower than that of polished flat glass. Pilkington's technique requires an experienced glass manufacturer and a large market to exploit the technology effectively. This firm has established full ownership or joint ventures, mostly with British Commonwealth countries of sufficient size to justify a local plant. In Europe, the United States, and Japan, the existence of established flat-glass manufacturers has led to licensing arrangements. Pilkington's policy is to cross-license subsequent patents and know-how without further payments between licensor and licensee as a means for encouraging the more rapid advancement of the industry as a whole.

International firms sometimes follow different principles in different product lines. Sulzer Brothers, a large Swiss engineering company, developed decades ago a marine diesel engine now used extensively in maritime fleets throughout the world. These engines are manufactured in ship-building countries by marine engineering firms under license from Sulzer. More recently, after a substantial research and development effort, they successfully introduced a shuttleless loom into the world market. Ten years after the start of commercial production, they still manufacture these looms in their Swiss factory for export to world

markets through their own sales outlets. As long as Sulzer is technically ahead, they delay transferring manufacturing know-how. When transfer is deemed advantageous, they prefer majority participation to assure the maintenance of technical standards and to maximize earnings; minority shares are accepted only where it enhances market entry. Their overall policy is to avoid granting exclusive licenses, so that they may continue to sell in that market should the licensee fail.

Concluding Remarks

The international corporation has emerged as an important instrumentality of resource allocation in the world economy. This role goes well beyond the classical concept of market mechanisms and is more in the mainstream of Schumpeter's contribution on the entrepreneurial role. Major factors influencing pricing and allocation efficiencies are how the contracting parties view the costs and benefits of technology transfer and the competitiveness in donor and recipient markets. It does not necessarily follow that an intensification of technological dissemination will lead to more effective utilization of world resources; transfer induced by protection of national markets increases overall resource costs of world production and nurtures technological stagnation. But flexibility on the part of international corporations in choosing alternative modes of transfer can lead to a wider dissemination of technology and better resource utilization throughout the world.

Between industrially-advanced partners, open-licensing policies, which includes cross-licensing, contribute to technological progress. How to induce these practices is a matter of public policy, beyond the scope of this paper. Similarly, public policies to assure competitive pricing of technology relates to larger issues of competition in sup-

plier markets and the relative bargaining position of purchasers. Issues of excessive rent to technology donors, including returns to indivisible packages of production and marketing services, are also part of larger questions of market structure and bargaining positions.

Developing countries are in a somewhat disadvantageous position in acquiring technology at competitive costs. This stems from a dearth of financial resources and basic deficiencies in both technological absorptive capacity and in the broader spectrum of production and marketing capabilities. An even more fundamental issue for the developing countries is whether they can acquire needed technology through licensing arrangements, and thereby avoid direct investment and associated foreign control. National policies may be undertaken along the lines of measures adopted in Japan to reinforce the bargaining position of local firms and their access to technology sellers in the world market.¹² More complete knowledge of what the world market has to offer in the way of technology suppliers would improve procurement efficiencies. The fact is that certain international firms are more willing and able than others to adjust industrial transfer to the specialized needs of developing economies.¹³

¹² See author's "Role of Science and Technology in Advancing Development of Newly Industrializing States," *Socio-Economic Planning Sciences*, Dec., 1969.

¹³ For example, DuPont of Mexico has been active in adapting acquired technology to the scale of demand of the Mexican market without incurring an excessive cost premium. Cummin's Engine Company has designed and built a diesel series to meet market demands both in developing countries and in certain specialized markets of industrialized areas. Massey-Ferguson in Mexico and Fiat in Yugoslavia have both structured their investments to minimize diseconomies of small-scale production and built-in technological obsolescence. See author's *Industrial Technologies for Developing Economies* (Praeger, 1969), pp. 45-53.

THE EFFICIENCY (CONTRADICTIONS) OF MULTINATIONAL CORPORATIONS

By STEPHEN HYMER
Yale University

Multinational corporations are a substitute for the market as a method of organizing international exchange. They are "... islands of conscious power in an ocean of unconscious cooperation," to use D. H. Robertson's phrase.¹ This essay examines some of the contradictions of this latest stage in the development of private business enterprise.

At the outset, we should note that the multinational corporation raises more questions than economic theory can answer. Multinational corporations are typically large firms operating in imperfect markets and the question of their efficiency is a question of the efficiency of oligopolistic decision making, an area where much of welfare economics breaks down, especially the proposition that competition allocates resources efficiently and that there is a harmony between private profit maximization and the general interest. Moreover, multinational corporations bring into high definition such social and political problems as want creation, alienation, domination, and the relationship or interface between corporations and national states (including the question of imperialism), which cannot be analyzed in purely "economic" terms.

I. *Division of Labor and the Extent of the Firm*

Our starting point is the fact that there are two kinds of division of labor: the division of labor between firms coordinated by the markets; and the division of labor within firms, coordinated by entrepreneurs. International trade theory has been mainly concerned with the first of these and has long stressed the desirability of widening international markets to increase the division of labor and exchange. Far less attention has been paid to the parallel proposition that the division of labor within a firm is limited by the extent of the firm and the economic and social questions this raises.

Unfortunately, the empirical evidence is not very helpful in deciding the degree to which large international firms should be encouraged in order

to reap the benefits of internal specialization and exchange. Few studies have been made on the relationship of foreign investment to a firm's overall efficiency and as far as quantitative evidence is concerned, we must view this question as a completely open one. With regard to the effect of size, the evidence is more plentiful but not conclusive. A number of studies on differences in performance of large and small firms have in general concluded that firms experience economies of scale up to a certain minimum size, after which there is little relationship between size and performance. Applying these results to the multinational corporation suggests that most parent firms are large enough to have exhausted economies of scale without foreign investment, although many of their subsidiaries may be too small to stand on their own feet.

These tests, however, have several inadequacies and may seriously underestimate the advantages of size. The major difficulty is that large firms are seldom engaged in exactly the same activities as medium-sized or smaller firms and their performance is not really comparable. The fact that very large firms do not seem to be significantly more profitable than their smaller rivals or to grow significantly faster does not preclude the possibility that they are specializing in activities where size is of great advantage and which would not be undertaken if the large firms did not exist. The structure of output within a country could well be a function of the size distribution of its firms without there being observable differences between large and small firms with regard to the more commonly studied characteristics.

The qualitative evidence on the structure of business enterprise and its evolution through time suggests that both size and internationality have important positive effects on a firm's strength and ability. Since the beginning of the industrial revolution there has been a steady increase in the size of manufacturing firms, so persistent that it might almost be formulated as a general law of capital accumulation. These increases in size were accompanied by important changes in organizational structure involving both increased subdivision or differentiation of tasks and increased integration through the creation of new organs of control. Business administration became a highly

¹ D. H. Robertson quoted in R. H. Coase, "The Nature of the Firm," *Economica*, New Series, 1937, pp. 386-405. Reprinted in G. S. Stigler and K. E. Boulding, *Readings in Price Theory* (Richard D. Irwin, Inc., 1932).

specialized activity with its own elaborate division of labor; and the corporation developed a brain to consciously coordinate the various specialties and to plan for the survival of the organism as a whole.

Chandler² distinguishes three major stages in the development of corporate capital. First, the Marshallian firm, organized at the factory level, confined to a single function and a single industry, and tightly controlled by one or a few men who, as it were, see everything, and decide everything. The second stage emerged in the United States at the end of the nineteenth century when rapid growth and the merger movement led to large national corporations, and a new structure of administration was developed to deal with the new strategy of continent-wide, vertically integrated production and marketing. The family firm gave way to the modern corporation with a highly elaborate administrative structure to organize the many disparate units of a giant enterprise. The next stage, the multidivisional corporation, began in the 1920's and gathered great momentum after the second World War. It too was a response to a new marketing strategy. To meet the conditions of continuous innovation, corporations were decentralized into several divisions, each specializing in one product line and organized as an almost autonomous unit similar in structure to the national corporation. At the same time, an enlarged corporate brain was created in the form of the general office to coordinate the various divisions and to plan overall growth and survival. This form is highly flexible and can operate in several industries and adjust quickly to rapidly changing demands and technology.

With each step in the development of business administration, capital obtained new power and new horizons. As Chandler and Redlich³ point out, there are three levels of business administration. Level three, the lowest level, is concerned with managing the day-to-day operations of the enterprise; i.e., keeping it going within the established framework. Level two is responsible for coordinating the managers at level three. Level one's function is goal determination and planning; i.e., setting the framework for the lower levels. In the Marshallian firm all three levels are embodied in one entrepreneur. In the national corporation,

the top two levels are separated from the bottom level. In the multidivisional corporation, differentiation is far more complete; level three is completely split off from level two and is concentrated in the general office whose specific function is strategy, not tactics.

In other words, the process of capital accumulation has become more and more specialized through time. As the corporation evolved, it developed an elaborate system of internal division of labor, able to absorb and apply both the physical sciences and the social sciences to business activity on a scale which could not be imagined in earlier years. At the same time, it developed a higher brain to command its very large concentration of wealth. This gave it the power to invest on a much larger scale and with a much wider time-horizon than the smaller, less developed firms that preceded it. The modern multidivisional corporation is thus a far cry from the Marshallian firm in both its vision and its strength. The Marshallian capitalist ruled his factory from an office on the second floor. At the turn of the century, the president of a large national corporation was lodged in a higher building, say on the seventh floor, with wider perspectives and greater power. In the giant corporation of today, managers rule from the top of skyscrapers; on a clear day, they can almost see the world.

Each step in the evolution of business enterprise had important implications for the structure of the international economy, just as each excursion into the international economy provided new challenges to the corporation and speeded its evolutionary development. In a world of Marshallian firms, commodity trade and portfolio capital were the main engines of international exchange. Movement of enterprise between countries was sharply limited because firms were small and lacked the appropriate administrative structure. The diffusion of Marshall's vital fourth factor, organization, from advanced to less advanced countries was therefore exceedingly slow. Movements of portfolio capital were substantial, at times, because the small Marshallian firms were associated with a highly developed banking and financial system. But the ability of less advanced countries to absorb capital (and technology) was limited to the rate at which they could build up their own organizations, a slow and difficult process given the negative policies of most governments in Africa, Asia, and Latin America, especially those in colonial dependencies. The range of goods which could be produced was thus restricted and the possibility for international

² Alfred D. Chandler, *Strategy and Structure* (Doubleday & Co., 1961).

³ Alfred D. Chandler and Fritz Redlich, "Recent Developments in American Business Administration and Their Conceptualization," *Bus. Hist. Rev.*, Spring, 1961.

trade to equalize factor prices was severely limited.⁴

The national corporation opened new possibilities of transferring organizational abilities internationally. The new administrative structure and financial power enabled firms to undertake direct foreign investments and organize large-scale production in mining and manufacturing in foreign countries. However, this migration of business enterprise occurred only on a limited scale and was usually restricted to a narrow activity; i.e., to acquiring raw materials used by the parent company or to exploiting some technological advance or differentiated product developed by the parent company. Moreover, to the extent that investment strengthened the firm's market control, its effect was considerably less beneficial and perhaps even negative.

The modern multidivisional or conglomerate enterprise is a much more powerful organizational form than the national corporation and appears capable of integrating world production and exchange to a much larger extent. Larger size and a more advanced administrative structure give it a much wider horizon leading in many cases to a global outlook and a transformation to the stage of multinational enterprise. It seems that after a certain point, a corporation comes to think in terms of its world market position rather than merely its United States or European market position and to plan in terms of worldwide factor availabilities and demand patterns. Since the process is just beginning, it is difficult to evaluate how strong this tendency will be. However, it is clear that at present large corporations are consciously moving towards an international perspective much faster than other institutions and especially much faster than governments, and are in the vanguard of planners of the new international economy created by the aeronautical and electronic revolutions. Since multinational corporations also have great financial and technical resources, they will certainly have many successes and will be able to speed up the spread of technology and to organize activities until now impossible. They are a large step forward but this is not, however, the same thing as saying that they serve the general interest as well as their own, that they are the best way to exploit the possibilities of modern science, or that they do not create

certain highly intractable problems which greatly impede their efficiency. We turn to these considerations.

II. *Bigness and Fewness*

Multinational corporations enlarge the domain of centrally planned world production and decrease the domain of decentralized market-directed specialization and exchange. Bigness is thus paid for, in part, by fewness, and a decline in competition since the size of the market is limited by the size of the firm. The precise effect of the present wave of direct investment on seller concentration in world markets is not well established. On the one hand, improved communications are breaking down barriers to trade and widening the market facing most buyers. On the other hand, direct foreign investment tends to reduce the number of alternatives facing sellers and to stay the forces of international competition. A great deal of statistical work needs to be done to evaluate the net effect of these two tendencies and establish the exact trend in the level of seller concentration, taking into account the growing international nature of the market. All that can be said at present is that the world level of concentration is much higher than it would be if foreign investment and domestic mergers were restricted. Since most countries are encouraging mergers at home and foreign investment abroad, for better or worse, the opportunity to increase competition by maintaining numbers is not being taken up.

Direct foreign investment thus has a dual nature. It is an instrument which allows business firms to transfer capital, technology, and organizational skill from one country to another. It is also an instrument for restraining competition between firms of different nations. Analyzing any particular case is an exceedingly complex matter, as the antitrust literature shows.⁵ For present purposes, the important point is to note that the general presumption of international trade economists in favor of free trade and free factor movements, on the grounds of allocative efficiency, does not apply to direct foreign investment because of the anticompetitive effect inherently associated with it. Just as in antitrust theory there are recognized reasons, within the frame-

⁴ Stephen Hymer and Stephen Resnick, "International Trade and Uneven Development," in J. N. Bhagwati, R. W. Jones, R. A. Mundell, Jaroslav Vanek, eds., *Kindleberger Festschrift* (M.I.T. Press, forthcoming).

⁵ See also Stephen Hymer, "Direct Foreign Investment and the National Economic Interest," Peter Russel, ed., *Nationalism in Canada* (Toronto: McGraw-Hill of Canada, 1966); Yale Economic Growth Center, Paper No. 108; "L'Impact des Firmes Internationales," in M. Bye, ed., *La Politique Industrielle de L'Europe Intégrée* (Paris: Presses Universitaires de France, 1968).

work of neoclassical economics, for preventing a firm from merging with another firm or from increasing its share of the market by growth, there are also international antitrust reasons for preventing a firm of one country from taking over a firm in another country or from acquiring or increasing its share of foreign production. Since this point can be easily misunderstood, it is important to stress that this is not a second-best argument but a genuine argument on antimonopoly grounds for interfering in international markets. A restriction on direct investment or a policy to break up a multinational corporation may be in some cases the only way of establishing a higher degree of competition in that industry. National antitrust measures cannot substitute for international antitrust when, for example, one of the major potential competitors to a domestic firm is its sister or parent affiliate within the same multinational group. In short, when we leave the conditions of perfect competition we lose the assumption of the invisible hand.

This argument, it should be noted, provides an important rationale for the infant entrepreneur argument supporting protection. Temporary protection of a weak firm from a stronger firm can improve the competitive structure of the industry in future periods by maintaining numbers. In the present context, the cost of this protection would have to be borne by the country that offers it while the benefits would accrue to the world as a whole. Thus, in reverse of the usual arguments, myopic behavior will lead to too little protection rather than too much. This presents a particularly acute problem in the case of underdeveloped countries. These countries typically do not sell commodities or buy capital or technology in competitive markets where there is an established price at which they can trade whatever quantity they want. Instead, they frequently face only a few potential buyers of their raw materials or their manufactured goods and a few potential sellers of a particular technology. The price they receive or pay therefore depends on their skill and strength in bargaining and not on market conditions alone. The less developed the country, the greater its disadvantage in the bargaining process because it has fewer organizations that are in any way a match for the giant companies with which it is dealing. Given the oligopolistic front maintained by the firms from developed countries, the underdeveloped countries need to devote an important share of their scarce resources to building up national enterprises which they can control and use in bargaining with foreign oligopolists. Ironically, their stronger bargaining position, by in-

creasing competitiveness, may improve general welfare in the rich countries as well—although it will harm those in the monopoly position.

III. *The International "Trickle Down"*

Many economists, in dealing with oligopoly, prefer to stress, as Schumpeter did, that the competition that counts lies in creative destruction through the introduction of new technology and new products. In that case, an oligopolistic market structure, even though it interferes with static optimum allocation, may be a necessary or at least a contributing factor to dynamic optimum allocation in a private enterprise system, because it allows innovators to capture some of the benefits of their discoveries and thus provides the incentive for research and development. The record of the United States shows that one certainly cannot fault oligopoly on the grounds that it does not produce a very rapid rate of technological change and product innovation. (Indeed it is easier to argue that the rate of change is too high.) One can expect international oligopoly via multinational corporation to provide the same kind of dynamic environment for the world economy as a whole.

The question of efficiency therefore hinges on the direction of change rather than the rate of change. An analysis of this problem involves an excursion into unexplored terrain since we do not now have an adequate theory on how corporations choose between the available paths of innovation. We certainly cannot assume that market forces compel firms to choose the optimum path. It is true that an innovation must, to some extent, meet the market test for a corporation to survive. However, what is at stake here is not whether the consumer has some choice but rather whether an oligopolistically competitive market structure provides him with the full range of choices possible.⁶ Oligopolists tend to copy each other, and their predictions as to what the consumer wants are often self-fulfilling, since in fact this is all that the consumer is offered. If we had only large numbers of independent decision centers could we assume that all avenues had been explored.

Since we cannot possibly treat this complex topic in any detail in the present paper, let us simply examine one theory of innovation closely associated with the multinational corporation and the international demonstration effect. The mark-

⁶ Sean Gervasi, "Publicité et Croissance Economique," *Economie et Humanisme*, Nov.-Déc., 1964. Harry Johnson, "The Political Economy of Opulence," in *The Canadian Quandary* (McGraw-Hill, 1962).

eting literature suggests new products typically follow a cycle known as trickle-down or two-stage marketing. An innovation is first adopted by a small group of individuals who act as opinion leaders and is then copied by others via the demonstration effect. In this process, the rich get more votes than everyone else, first of all because they have more money, second of all because they have discretionary income and can afford to be experimental, and, third, because they have high status and are more likely to be copied. The principle of consumer sovereignty cannot easily be applied to this process since, at most, only the special group in the first stage of the marketing process has something approaching a free choice. The rest have only the choice between conforming or being isolated.

In the international economy, trickle-down marketing takes the form of the international demonstration effect. Products are first introduced in the United States or Europe and then spread to other countries. Multinational corporations speed up this process by making it easier to transfer new products and marketing methods to less advanced countries. One of the key motives for direct investment, cited by corporations, is to gain control over marketing facilities in order to facilitate the spread of their products. If firms were denied control over communication and marketing facilities in the foreign countries and we had a regime of national firms (private or socialized) rather than multinational firms, the pattern of output would almost certainly be quite different than the one that is now observed. There would be more centers of innovation, and probably more variety of choices offered to the consumers, as each country developed products suited to its particular characteristics. Products from one country would spread to other countries either through trade or imitation but the movement would be coordinated by market competition rather than the planning decisions of top management in a few corporations whose interest it is to foreclose competition, to restrict the choices offered, and to insure the survival of their own organizations. It is difficult to speak with professional certainty in this badly neglected field, but it does not appear to be socially efficient to allow corporations to monopolize information on new possibilities created by science.

IV. *The International Hierarchy of Decision Making*

Marshall, like Marx, thought that the "chief fact in the form of modern civilization, the kernel

of the modern economic problem . . ."⁷ was the division of labor within the factory between those who plan and organize economic activity and those who work for them. In the modern corporation the hierarchical structure of command and authority has been greatly elaborated from the simple division between owners and workers in the Marshallian firm, but the tensions and conflicts of autocracy remain. They take on particular importance in the multinational corporation where problems of nationalism and problems of authoritarianism intertwine.

Multinational corporations are torn in two directions. On the one hand, they must adapt to local circumstances in each country. This calls for decentralized decision making. On the other hand, they must coordinate their activities in various parts of the world and stimulate the flow of ideas from one part of their empire to another. This calls for centralized controls. They must therefore develop an organizational structure to balance the need to coordinate and integrate operations with the need to adapt to a patchwork quilt of languages, laws, and customs. One solution is division of labor based on nationality. Day-to-day management in each country is left to nationals of that country who are intimately familiar with local conditions and practices and best suited to deal with local problems and local government. These nationals remain rooted in one spot, but above them is a layer of people who move around from country to country, as bees among flowers, transmitting information from one subsidiary to another and from the lower levels to the general office at the apex of the corporate structure. In the nature of things, these people, for the most part, will be citizens of the country of the parent corporation, just as we now find that the top executives of most of the major corporations in the United States are drawn from a relatively small homogeneous cultural group quite distinct from the population of the United States as a whole.

This creates two types of problems. In the first place, there is the internal problem of creating incentives for foreigners whose access to the top corporate positions will be necessarily limited. The second problem is far more important and is in the nature of an external diseconomy. The subsidiaries of multinational corporations are frequently amongst the largest corporations in their country of operations and their top executives play an influential role in the political, social, and cultural life of the country. Yet these people,

⁷ Alfred Marshall, *Principles of Economics* (Macmillan, 1961), pp. 74-75.

whatever their title, occupy at best a medium position in the corporate structure and are restricted in authority and horizons to a lower level of decision making. The country whose economy is dominated by the foreign investment can easily develop a branch plant outlook, not only with reference to economic matters, but throughout the range of governmental and educational decision making.

Thus there are important social and political costs to international specialization in entrepreneurship based on multinational corporations. The multinational corporation tends to create a world in its own image by creating a division of labor between countries that corresponds to the division of labor between various levels of the corporate hierarchy. It will tend to centralize high-level decision-making occupations in a few key cities in the advanced countries (surrounded by regional subcapitals) and confine the rest of the world to lower levels of activity and income; i.e., to the status of provincial capitals, towns, and villages in a New Imperial System. Income, status, authority, and consumption patterns will radiate out from the centers in a declining fashion and the hinterland will be denied independence and equality.⁸

This pattern contrasts quite sharply with the free trade system which offered both income equality and national independence. According to the factor price equalization theorem, trade allows a country to choose its own style and still share fully in the riches of the world. Whether large or small and even if its resource endowment is highly skewed, it can achieve factor price equalization with the rest of the world by varying the composition of output without surrendering its control over its capital stock and without the need for its members to leave the country to find employment elsewhere. Now the stakes seem to have gone up. In order to reap the gains from international exchange, a country has to become integrated into a corporate international structure of centralized planning and control in which it plays a very dependent role.

Countries may not be willing to play this game nor to completely break with it and the possibility arises, in part suggested by the Canadian experience, of getting the worst of both worlds. Canada has allowed an almost unrestricted inflow of capital and as a result has surrendered a great deal of national independence. At the same time,

she has adopted a number of policies, including high tariffs, which prevent international corporations from fully rationalizing production on a continent-wide basis. The record shows that foreign subsidiaries in Canada tend to perform at levels equal to their Canadian counterparts rather than at the higher levels of efficiency of their parent corporations. This suggests that many of the benefits of foreign investment have been emasculated while many of the costs remain.

V. *Big Corporations: Small Countries*

The efficiency with which multinational corporations can allocate resources internationally depends in large part on government policy decisions. If government decision making were independent of the structure of the private sector, we could view it as an exogenous factor and safely ignore it in an essay devoted to the multinational corporation. However, an increase in the importance of multinational corporations relative to national corporations will clearly have an important impact on both the ability and willingness of governments to carry out certain types of economic policies. An analysis of the efficiency of multinational corporations must take this into account and analyze, for example, its effect on government capital formation in the crucial sectors of infrastructures and human capital. This aspect is particularly important with regard to the problem of underdevelopment—clearly the greatest instance of inefficiency in today's international economy.⁹

Analyses of the role of foreign investment in underdeveloped countries often focus on the great disparity between the bargaining power of the corporation and the bargaining power of the government. The corporations are large and modern and have international horizons. The governments are typically administratively weak and have very limited information outside their narrow confines. In any particular negotiation between one country and one company, power in the form of flexibility, knowledge, and liquidity is usually greater on the private side than on the public side of the table.

The problem of unequal bargaining power can

⁸ This point is developed more fully in S. Hymer, "The Multinational Corporation and Uneven Development," in J. Bhagwati, ed., *Economics and World Order* (World Law Fund, 1970).

⁹ For a stimulating analysis of the relationship of multinational corporations to economic development, see: G. Arrighi, "International Corporations, Labour Aristocracies and Economic Development in Tropical Africa," D. Horowitz, ed., *The Corporations and the Cold War* (London, forthcoming); N. Girvan, "Regional Integration vs. Vertical Integration in the Utilization of Caribbean Bauxite," Lewis and Matthew, eds., *Caribbean Integration* (Inst. of Caribbean Studies, Univ. of Puerto Rico, 1967).

be illustrated with a simple model (developed in collaboration with Stephen Resnick).¹⁰ This model focuses on the feedback relationship between the government and the foreign corporation. The government provides certain support services to the corporation: protection, infrastructure, help in the creation of a labor force, land laws, etc. The corporation in return pays the government taxes and royalties. This is a trading relationship in which two main variables are involved: (1) the tax rate (t); and (2) the fraction of government expenditure devoted to support services (g). The outcome is determined by a process of bargaining which, for simplicity, can be viewed in a purely duopolistic form—one government and one country—though it, in fact, usually arises in a more complicated structure where there are several corporations and several power groups involved. The government, we assume, is interested in maximizing its surplus (total revenue from foreign firms less the cost of support services). The corporation is interested in maximizing profits after taxes. At one extreme the government may be very strong and choose (g) and (t) such as to make profit zero (we assume normal profits are included in cost) and to make the government surplus as large as possible. This seldom, if ever, occurs in underdeveloped countries where the bargaining tends to go in the opposite direction. The corporation sets (t) as low as possible, subject to the constraint that the government has enough money to: (a) provide necessary infrastructures; (b) remain in power and maintain law and order for the corporation. Since the government has little surplus it does not have the money to provide capital or services for other industries. This is in keeping with the foreign investor's interest, since the growth of other industries would compete away factors of production and would create interest groups who might challenge the corporation's hegemony. Provided that the political forces are kept under control in this system, the country can remain in its state of underdevelopment for a long period.

Such extreme cases are no longer possible because of the increased political strength of the local middle class in most underdeveloped countries and because of the changed nature of foreign in-

vestment. Modern multinational corporations are interested in manufacturing in underdeveloped countries and not just in raw materials and therefore want a growing market for advanced products and an educated, urbanized labor force. They are no longer tied to traditional backward governments, but have a stake in an active government sector which promotes growth and provides education and infrastructure. The "new foreign investment" is, then, a far cry from the "Banana Republic" kind, but important dangers remain. Statistics on income distribution show that the top one-third of the population typically gets about 60 percent of the total income. It is this top group which provides the direct and indirect labor force for large-scale manufacturing as well as the market. An alliance between this group and foreign investors represents a formidable bargaining force vis-à-vis the remaining two-thirds of the population. A government expenditure policy based on such as alliance would concentrate on the modern high-income sector, leaving the rest of the population as a source of unlimited supply of cheap labor for services and for menial work. Growth in these circumstances would retain its uneven quality and all the inefficiency that implies, albeit in a more advanced and progressive form than characterized the enclave economies of the previous round of foreign investment.

VI. *Multinational Corporations and Supranationality*

Multinational corporations create serious problems in the developed world as well. The most important of these, from the limited perspective of this essay, is that they reduce the ability of the government to control the economy. Multinational corporations, because of their size and international connections, have a certain flexibility for escaping regulations imposed in one country. The nature and effectiveness of traditional policy instruments—monetary policy, fiscal policy, anti-trust policy, taxation policy, wage and income policy—change when important segments of the economy are foreign-owned. This has long been recognized in countries such as Canada, but it is now becoming obvious that even the United States has reached the point where the international commitments of its corporations reduce the room for flexibility in national economic policy formation. If foreign investment continues to grow at anything like the rate of the last ten or fifteen years, this problem will become an extremely serious one for all North Atlantic countries.¹¹

¹⁰ S. Hymer and S. Resnick, "Interactions between the Government and the Private Sector in Underdeveloped Countries: Government Expenditure Policy and the Reflection Ratio," Ian Stewart, ed., *Economic Development and Structural Change* (Edinburgh: Edinburgh Univ. Press, 1969). Published in French as "Les Interactions entre le Gouvernement et leur Secteur Privé," *L'Actualité Économique*, Oct.-Dec., 1968.

¹¹ For an attempt to predict the trend towards mul-

This contradiction between multinational corporations and nation states has important bearing on the efficiency of the multinational corporation. The main problem, stated most simply, is as follows: if national power is eroded, who is to perform the government's functions? For example, if nation states, because of the openness of their economy, cannot control the level of aggregate economic activity through traditional monetary and fiscal policy instruments, multinational agencies will need to be developed to maintain full employment and price stability. Yet such organizations do not exist at present, nor can they be quickly built. Either one must argue that the Keynesian problem has somehow been solved by the creation of the multinational corporation (along with a host of other problems) or else one must agree that it is not feasible to have international business integration via direct foreign investment proceeding at a much faster rate than political integration. Yet, this seems to be precisely what is happening. Most of the large American firms have already staked out their claims in the European market and many of the leading European firms are now rapidly entering foreign markets, including those of the United States. A predominance of multinational corporations in the North Atlantic economy seems therefore to be a *fait accompli*. Government cooperation is not growing at anywhere as rapid a rate. If serious problems arise, governments are likely to reassert their power and attempt greater regulation and control over the business enterprises within their jurisdictions. Economists will rightly point out that these restrictions create inefficiencies in the allocation of the economic resources. It is important, however, to realize the role played by a too

tinationalism and the problems it causes, see Stephen Hymer and Robert Rowthorn, "Multinational Corporations and International Oligopoly: the Non-American Challenge," in C. P. Kindleberger, ed., *The International Corporation* (M.I.T. Press, 1970).

liberal policy towards private capital, movements and mergers that created the multinational industrial structure.

VII. Conclusion: Some Subjective Evaluations

This essay has presented a list of advantages and disadvantages of multinational corporations. Assuming there are no important omissions and that each point taken by itself is valid, the question arises as to what weights should be attached to the various arguments. One simple summation, offered here without proof, is as follows: The large corporation illustrates how real and important are the advantages of large-scale planning, but it does not tell us how best to achieve wider domains of conscious coordination. Broadly speaking, there are two main directions in which one can proceed. Multinational corporations integrate one industry over many countries. The alternative is to integrate many industries over one country and to develop noncorporate linkages between countries for the free flow of goods and, more important, the free flow of information. The advantage of the second direction is that it keeps the economy within the boundary of the polity and the society. It thus causes less tension and creates the possibility of bringing economic power under control by removing the wastes of oligopolistic anarchy. This would allow more scope for solving the two major economic problems of today, affluence and poverty, than the first alternative. The trend, however, is clearly in the direction of the first alternative. The coming age of multinational corporations should represent a great step forward in the efficiency with which the world uses its economic resources, but it will create grave social and political problems and will be very uneven in exploiting and distributing the benefits of modern science and technology. In a word, the multinational corporation reveals the power of size and the danger of leaving it uncontrolled.

DISCUSSION

THOMAS D. WILLETT: In *The Structure of Scientific Revolutions*, Thomas Kuhn discussed the progress of science in terms of the movement from one paradigm or view of the world to another. Kuhn described normal science as a cumulative progression of puzzle solving within the context of a generally shared theoretical framework or approach. Confrontation with an increasing number of anomalies is not in itself sufficient to cause rejection of a paradigm, but such anomalies may create a period of crisis in which new theories or approaches are generated and compete with the initial paradigm for widespread acceptance, often with the result that one of these alternative approaches attains the status of a new paradigm.

International trade theory gives evidence that it currently is in the midst of such a revolution.¹ In the postwar period trade theory has been dominated by the two-factor version of the Heckscher-Ohlin model of comparative advantage. The perfection of this model, or what Kuhn would term puzzle solving within this paradigm, has until recently been unquestionably the major activity of trade theorists. But when confronted with evidence concerning actual trade patterns, the implications of the two-factor Heckscher-Ohlin model do not stand up well. Prodded by Leontief's paradox and subsequent empirical work, more attention has been paid to an increasing number of factors of production such as natural resources and the many various forms of capital such as human skills and information and technology and social overhead capital. The increasing recognition (or in some cases the rediscovery) of the importance of economies of scale, transportation costs, demand patterns, and product differentiation have likewise contributed to the beginnings of a new paradigm which Harry Johnson referred to in his recent Wickseil lectures as an emerging dynamic theory of international trade and factor movements.² This revolution in international trade theory has not reached the general level of rigor of its predecessor, but it is an approach in which I find much richer possibilities.

One of the effects of this newly emerging approach has been increased interest in trade and factor move-

ments as alternative methods of servicing markets and the role of the international corporation in this process. Hence the papers prepared for this session on the international firm and efficient economic allocation are very much within the spirit of this new approach to international trade theory and make, I feel, a very useful contribution to its advancement.

My primary assignment is to comment on Professor Baldwin's excellent paper on trade in inputs and outputs. I am afraid that I have little criticism to offer on this careful, concise survey. Let me raise just one question concerning Baldwin's argument that while a high degree of risk and uncertainty acts to discourage all types of input and output flows, this condition particularly impedes capital movements due to the indivisibility and comparative immobility of most specific capital goods in which foreign capital funds usually become embodied. I think that on this question it is important to distinguish between types of risk and uncertainty. Where the major source of such concern is over the possible imposition of controls, I would agree with Professor Baldwin's conclusion; but where the major international risk is the possibility of exchange rate fluctuations, this conclusion may be reversed. In his paper prepared for the Bürgenstock Conference on Exchange-Rate Flexibility, John Watts has cogently argued that greater exchange rate fluctuations would have a much greater impact on trading (export-import) firms than on the large multinational corporations.

Let me now turn to Professor Hymer's paper. I am bothered by several aspects of Hymer's discussion. Let me briefly mention two. One is the recurrent theme that while trade is competitive, direct investment is anticompetitive. Hymer views direct foreign investment as "an instrument for restraining competition between firms of different nations" and argues that while "on the one hand improved communications are breaking down barriers to trade and widening the market facing most buyers. On the other hand, direct foreign investment tends to reduce the number of alternatives facing sellers and to stay the forces of international competition." I do not think that the validity of this last statement is at all clear. I very much agree with Tibor Scitovsky that competition should not be judged by the number of producers coexisting in a market, but rather the ability and willingness of the individual firm to encroach upon the markets of others (*Economic Theory and Western Economic Integration*, p. 124). Of course judgments on competitiveness depend on what one views as the most relevant alternative. As viewed against the benchmark of classical perfect competition, the large multinational corporations may not score well. But viewed against a benchmark such as the market structures existing in Europe before the large-scale invasion of U.S. subsidiaries, there is considerable evidence of increased competition as a result of the multinational corporation. A foreign intruder may be less likely to comply with

¹ For discussions of the development of economic thought from the point of view of Kuhn's approach, see A. W. Coats, "Is There a 'Structure of Scientific Revolution' in Economics," *Kyklos*, Fasc. 2, 1969; and Donald Gordon, "The Role of the History of Economic Thought in the Understanding of Modern Economic Theory," *A.E.R.*, May, 1965. As is indicated by Kuhn, scientific revolution or changes in paradigms need not affect an entire discipline (such as the Keynesian revolution, for instance), but may also occur within subdisciplines.

² A condensation of Johnson's Wickseil lectures along with discussion by Richard Caves and others appears in Paul A. Samuelson, ed., *International Economic Relations: Proceedings of the Third Congress of the International Economic Association* (St. Martin's Press, 1969).

explicit or implicit restrictive pricing agreements, and there is considerable evidence that competitive pressures from U.S. subsidiaries have stimulated modernization and productivity improvements in European industry, generating increases in what Harvey Leibenstein has termed "X-Efficiency." As Hymer has stressed in past writings, the major rationale for direct investment is the exploitation of some degree of monopoly power. But often this is a temporary monopoly resulting from the recent creation of new economically relevant knowledge which will either be superseded by subsequent discoveries or gradually find its way into the public domain. As Johnson emphasized in his Wickseil lectures, at the heart of the political and policy problems raised by the international operations of the modern corporation is the conflict between recognition that the cost of creating new knowledge has to be reimbursed if there is to be a private incentive for investment in knowledge creation and that such knowledge once created is a public good whose monopolization restricts its usage to less than a statically efficient level.

In closing, let me briefly supplement Hymer's discussion of the inequality of bargaining power between big corporations and small countries. I think that Hymer's general point is well taken, but I find it interesting to note that the rather extensive theoretical literature on capital movements and national advantage suggests that divergencies between private and social costs and benefits would generally lead to free market capital exports in excess of what would be optimal from a static nationalistic point of view.⁸ One strand of this argument is essentially the optimum tariff argument applied to capital movements. Another and probably the quantitatively most important divergence results from the prevalence of double tax agreements which allow foreign subsidiaries to count tax payments to the host country as a credit against their tax liabilities in their home country. While the corporation should be reasonably indifferent as to which government it pays taxes, national interests are not. Such tax arrangements give host countries an incentive to tax foreign investment at least up to the level of the capital exporting country. Of course, these tax revenues to the host country are usually not completely free, because the presence of the foreign subsidiary may entail increased government costs, and it is possible that special incentives given to foreign investment may in some instances exceed the social benefits generated. But in general double taxation arrangements would seem to go a long way to-

ward reducing such possibilities. It would also appear that where capital exporting countries are engaged in policies to aid capital receiving countries, the foregoing of possible national advantages from restricting capital exports would be a wise course of action.

STEFAN H. ROBOCK: Most of the current literature on the international business phenomenon accepts the idea, with little or no qualification, that technology transfers through the international firm contribute to more efficient economic allocation. The tendency seems to be to idealize the positive and overlook the negative.

The reality is that technology transfers through the international firm may or may not contribute to more efficient economic allocation. Furthermore, the test of efficiency will be different from the viewpoint of the international firm, the world as a whole, and the individual nation. To inject more balance into the picture, I would like to identify several types of situations where technology transfers through the international firm do not contribute to efficient economic allocation.

Dr. Baranson's interesting paper deals mainly with factors influencing the logistics of technology transfers and the conflict of interest in licensing versus investment decisions. But in his conclusions he does note that "it does not necessarily follow that an intensification of technological dissemination will lead to more effective utilization of world resources." A "transfer induced by protection of national markets," he notes, "increases overall resource costs of world production and nurtures technological stagnation."

As Baranson suggests, it may be a more efficient use of world resources to supply a foreign market through exports from an established production facility of economics scale than to transfer technology so that a local plant of uneconomic size can be established behind tariff or import control barriers. But even this apparently clear-cut case must be qualified. From a national rather than a world point of view and measured by national development goals and a long-term time horizon, the technology transfer may be contributing to a more efficient national allocation of resources.

The Baranson example focuses on inefficiencies due to uneconomic scale of production and protection from competition. Another type of situation involves what I call "technology fit." Here I am referring to differences among countries in the relative costs and supply of production factors. For example, industrial technology developed in the United States is strongly shaped by the large size of the U.S. market and the relatively high cost of American labor. When transferred to other countries, even after the international firm or the license makes minor adaptations in products or production techniques, the technology may not result in efficient economic allocation of local resources either from a world or a national point of view. This is particularly true for the less developed countries if alternative technology can be developed that is less capital intensive and more labor intensive and that can produce competitively on a smaller scale.

The problem of technology fit has already become a

⁸ See, for instance, J. Carter Murphy, "International Investment and the National Interest," *So. Econ. J.*, July, 1960; R. W. Jones, "International Capital Movements and the Theory of Tariffs and Trade," *Q.J.E.*, Feb., 1967; and W. M. Corden, "Protection and Foreign Investment," *Econ. Record*, June, 1967, and the references cited in these papers. Specifically on the bargaining question see Kotchi Hanada, "Strategic Aspects in Taxation of Foreign Investment Income," *Q.J.E.*, Aug., 1966.

matter of intense concern in the less developed countries where capital is scarce and manpower is plentiful. And there is scattered but persuasive evidence that different types of technology can be developed that are efficient and that better fit the situation in those countries. As a result, there have been suggestions that the less developed countries should join in a cooperative research and development effort to develop technology more appropriate to their needs than that being transferred through international firms. But the more widely the existing technology of the large and advanced countries is adopted, the more difficult and expensive it is to change to a better one.

The movement toward economic integration may provide an antidote to the problems of scale economies and absence of competition. But enlarged markets still leave factor proportions that are most economic in the country of birth frozen into the technology. Thus transferred technology may not maximize the comparative advantages of the environment into which it is transplanted. And I might add parenthetically that, given the discouraging record of integration attempts in the less developed countries, a cooperative research effort to enlarge technology options may be more feasible politically than successful economic integration.

Still another type of negative effect associated with technology transfers by international firms is the possibility that the growth of research and development activities in the receiving country will be stunted. This possibility is suggested as a general feature of foreign investment in a fascinating essay by Albert O. Hirschman entitled, "How to Divest in Latin America and Why."¹ Hirschman points out that technology transfers and other features of foreign investment at their creative best bring missing factors to the recipient countries. At a later stage, however, the international firm may play a stunting role in relation to the development aspirations of the country.

One reason, according to Hirschman, is that institutional inertia makes for importing of the missing factors even after a country reaches the point where it can supply such factors more efficiently locally than through imports. Another reason is that the managers of foreign-owned plants are frequently too restrained and inhibited in contributing an industrial voice to economic policy making of the host country in which their companies are "guests." This is particularly important when a developing country needs new industrialists able to speak with a strong and even militant voice in support of national policies such as creating an expanded research and development capacity.

The limits of time do not permit me to expand the list of negative potentials. But the situations already mentioned should demonstrate that technology transfers through the international firm are a mixed bag from the standpoint of efficient economic allocation. We need more sorting and more empirical investigation, guided by a more complex conceptual frame-

work, to know where the balance lies between less and more efficient economic allocation. We should recognize, however, that present patterns are not likely to persist. Most receiving countries have not yet developed sophisticated policies toward technology transfers. Where control has been exerted, receiving nations have often been most concerned about the prices charged to either licensees or affiliates of international firms for acquired technology. Increasingly, I would expect countries to follow the advice of a leading scholar in the field and "work toward national policies which are less defensive and more positively-oriented toward the acquisition and diffusion of technologies" that the international firms have to offer. To do this, "each country must develop a strategy attuned to its particular resources and problems."² The Japanese experience offers a successful example that is bound to stimulate and encourage other countries.

To summarize, the contribution to technology transfers to efficient economic allocation must be evaluated differently from the standpoint of the world as a whole, the receiving nations, and the international firm. The international firms tend to emphasize the positive features. The receiving countries emphasize the negative features even while continuing to encourage the expansion of international business. The challenge to scholars is to improve our conceptual framework and enlarge the body of factual data so that the phenomenon is better understood from an objective point of view.

RONALD I. MCKINNON: My assignment is to discuss Professor Hymer's provocative paper. Because his argument is diffuse in many places, I shall briefly reconstruct his main thesis below:

1. The nature of technical change has been such as to make the growth of large companies optimal within developed nation states.
2. These organizations have great competitive advantages in transferring technology efficiently to less developed countries by operating directly there.
3. However, there are contradictions and perversities in the transfer process which individual host governments cannot deal with adequately because of their limited jurisdictions and small size. Domestic monetary and fiscal policy can be easily subverted by international flows of capital and commodities. There may be undue exploitation by large corporations of small countries and domestic entrepreneurial development may be retarded.
4. There is, then, a need for a worldwide bureaucracy or policy towards multinational corporations which can exercise full legal control over them. Alternatively, access to the domestic markets of developing countries by the international corporation should be limited.

This summary cannot do full justice to Professor Hymer, but I shall take a position which is quite

¹ *Essays in International Finance*, No. 76, Nov., 1969 (Int. Fin. Sec., Princeton Univ.).

² James Brian Quinn, "Technology Transfer by Multinational Companies," *Harvard Bus. Rev.*, Nov.-Dec., 1969, p. 160.

different, not only from his, but possibly from the whole theme of this session. Although it is true that the multinational corporation has become a pervasive economic phenomenon, it does not follow that the multinational legal entity is of much analytical use in dealing with the main policy problems of economic development: taxation, foreign trade, capital markets, technology transfers, etc. Moreover, preoccupation with the need for international controls may be something of a red herring because it detracts from focusing on perverse policies which developing countries are now undertaking, and may even provide some intellectual rationale for such policies. In order to be quite specific on this point, I will emphasize how many countries have mismanaged their financial and foreign trade policies so as to retard their own development and so as to make the multinational corporation much more dominant than would otherwise be the case. There are, of course, areas other than finance and trade where policy has gone astray.

The Financial Problem. A well-developed internal capital market is a necessary condition for domestic entrepreneurship to flourish. Without sources of finance which are available to individual enterprise with high profit opportunities, local businessmen cannot get finance for manufacturing of either intermediate products or final consumer goods; and farmers cannot get finance to buy the package of new seeds, pesticides, fertilizers, etc., which is necessary to take advantage of the new agricultural technology. At the earlier stages of development, the banking-monetary system is necessarily the most important source of domestic finance and public policy towards it is of critical importance.

However, in a variety of countries in Latin America, in Asia, and in Africa the "organized" financial sector has been quite severely repressed. One common technique is to maintain a significant rate of inflation with fixed deposit rates of interest (and other usury laws) on various classes of bank deposits. The amount of real finance undertaken by the banking sector becomes very limited. Another is to maintain very high reserve requirements in such a way as to make the banking system little more than the financial arm of the government which is unavailable to the private sector. The traditional curb markets with small-scale moneylenders and pawnbrokers are certainly better than nothing as a source of private finance, but they are generally inadequate for industrial development or agricultural modernization. The domestic capital market remains highly fragmented with much capital put to inferior uses at home and abroad. The rate of return to be earned on Swiss bank accounts may itself be very low, but at least it is somewhat higher than can be obtained on liquid assets in the home country.

In the presence of this substantial capital flight, we have the paradox of the large international corporation coming into various underdeveloped countries and buying up its assets—sometimes at bargain-basement prices. Domestic entrepreneurs have no financial resources with which to compete successfully even when they have superior technical information.

Unsurprisingly, investment coming in in this form generates a lot of hostility and perverse regulations, political controls, etc. A rationalization of the financial structure would serve a dual purpose of limiting the advantages that foreigners have in the domestic market and at the same time increasing the average rate of return to new investments. Direct regulation of the multinational corporation is a third- or fourth-best alternative.

There is no inherent reason why financial rationalization cannot take place. Japanese development has occurred independently of foreign direct investment, in large measure because there were plentiful sources of domestic finance. The Japanese banking structure in both the industrial and agricultural areas was relatively highly developed early in Japan's industrialization, and the government has generally been a net contributor to the banking system as a source of finance; that is, the government has frequently tended to run surpluses on current account and the surpluses have been invested through the banking sector in a variety of private and quasi-governmental activities. With a highly developed domestic capital market, the Japanese could afford to limit direct foreign investment and to depend on licensing of foreign technologies by domestic firms. The fact that they could depend on domestically controlled firms in their development processes undoubtedly did accelerate the process of learning by doing.

The Foreign Trade Problem. Let me now turn to foreign trade policy which is necessarily closely linked with financial policy and has been equally perverse in attracting the wrong kinds of foreign investment. Professor Hymer has mentioned that one could make a case for "protection," the form of which he does not specify, in less developed countries to protect against the predatory international corporation and possibly to increase worldwide competition by developing new national bases for industrialization. Indeed Hymer goes so far as to suggest that we may have too little protection in developing economies because they are myopic with respect to world welfare, whereas new industrial bases under independent control would benefit the rich by diminishing the monopoly power of multinational corporations.

If one assumes that protection means tariffs and other restriction on commodity trade (which is indeed what protection has meant to actual policy carried out in poor countries), such restrictions on foreign trade have been associated with turning open economies into closed ones. Protection has had the general equilibrium impact of making both national entrepreneurs and multinational corporations operating in poor countries less competitive in world markets than they would otherwise be. Rather than increasing worldwide competition, trade restrictions have generally succeeded in reducing it. Many noncompetitive domestic cartels of small and large firms are maintained in place by the licensing-issuing authorities who derive their authority from foreign trade controls.

Besides closing economies off from foreign trade,

protection has also succeeded in setting up highly differentiated tariffs for various classes of domestic producers; that is, manufactured consumer goods and other items whose domestic production is readily identifiable generally receive very heavy protection. At the same time, capital goods and other intermediate products may enter duty-free and receive especially favored exchange rates. The relationship of the domestic exchange rate to the internal price level then gives strong incentives for all domestic companies, large and small, nationally-owned or foreign, to over-use imported intermediate inputs. We are all familiar with the phenomenon negative value added.

It should be emphasized, however, that this discriminatory tariff structure gives a particular operating advantage to foreign corporations, which naturally depend more heavily on imported capital goods and raw materials—perhaps produced elsewhere within the same corporation. Domestic entrepreneurs, who would more naturally turn to domestic sources of raw materials and capital goods, are in fact put at a competitive disadvantage by the highly favorable terms on which these imports can take place. In addition, potential local producers of raw materials and capital goods face overly severe competition from abroad and may never come into existence. Thus, efficient backward linkage has been cut off and local economies have participated less in new product development and technological innovation than might otherwise be the case.

One could go on and discuss how domestic fiscal policy and labor policy are also perverse, but space does not permit. Let me then restate my principal conclusion. Preoccupation with the multinational corporations simply focuses on the wrong problem and can even provide justification for following misguided internal policies which may even unduly accentuate the role of the multinational corporation in the domestic economy. There is nothing inevitable in the development process which requires poor countries to start with "the highest form of capitalism," but rather

there is plenty of evidence to show that very small companies that do receive adequate local financing can compete in international export markets, as the recent experience of Taiwan, Korea, and Hong Kong would seem to indicate. It is necessary, however, to have plentiful sources of internal domestic finance and a relatively free commodity trade policy so that exporting "looks" profitable. This "Japanese" route is still open to LDC's today where the possibilities for engaging in international commerce are even more favorable than in the early development of Japan.

As a postscript, I would comment on Professor Hymer's concern with the loss of internal monetary fiscal autonomy within advanced countries of Western Europe and within LDC's. In the postwar period, business cycles have largely been internal in origin; that is, national economic policies have broken down for one reason or another or certain shocks or upheavals have occurred on a national basis. One can look at the experience of Italy in 1962-63, France in 1968, Germany 1966-67, and the United States now as illustrations from the more advanced countries, but of course there are many more examples one can pull from LDC's. Throughout all this, the international economy has remained relatively strong and stable. It has been a considerable advantage to individual nation states that they remain hooked up to the international economy and can use variations in their trade balances to soak up what would otherwise be more significant domestic fluctuations. One could even go further and say that it is very fortunate that domestic policy has been subverted from time to time by international events; otherwise the policy would work out worse than it did. As an illustration, one can take the current efforts of the American banking system to subvert Regulation Q by manipulations in the Euro-dollar market. With our decline in confidence in "fine fiscal tuning," it seems that we are not giving up very much if we do indeed lose a certain amount of national monetary and fiscal autonomy because of the multinational corporation.

NEW IDEAS IN PURE THEORY*

PROBLEMS IN THE THEORY OF MARKETS UNDER UNCERTAINTY†

By ROY RADNER

*University of California, Berkeley
and Churchill College, Cambridge*

Introduction

One of the notable intellectual achievements of economic theory during the past twenty years has been the rigorous elaboration of the Walras-Pareto theory of value; that is, the theory of the existence and optimality of competitive equilibrium. Although many economists and mathematicians contributed to this development, the resulting edifice owes so much to the pioneering and influential work of Arrow and Debreu that in this paper I shall refer to it as the "Arrow-Debreu theory." (For a comprehensive treatment, together with references to previous work, see [6].)

The Arrow-Debreu theory was not originally put forward for the case of uncertainty, but an ingenious device introduced by Arrow [1], and further elaborated by Debreu [5], enabled the theory to be reinterpreted to cover the case of uncertainty about the availability of resources and about consumption and production possibilities. (See [6, Chap. 7] for a unified treatment of time and uncertainty.)

In the present paper I take the Arrow-Debreu theory as a starting point and discuss certain extensions, limitations, and possible new departures. In particular, I: (1) show how the theory can be extended to account explicitly for differences in information available to different economic agents, and for the "production" of information; (2) present a critique of the (extended) theory, especially its failure to explain or take account of money, stock markets, and the presence in the real world of active markets at every date; (3) argue for the consideration of a theory of a sequence of markets and suggest several concepts of equilibrium that might be appropriate to such a theory; and (4) present some results on the existence of an equilibrium of plans, prices, and price expectations in a sequence of markets.

* The papers by Herbert Scarf, "An Example of an Algorithm for Calculating General Equilibrium Prices," and by Lloyd S. Shapley and Martin Shubik, "On the Core of an Economic System with Externalities," were printed in the September, 1969, issue of the *A.E.R.* but presented at this session.

† This paper is based on research supported in part by the National Science Foundation.

The main features of the Arrow-Debreu theory have been available in the literature for more than a decade and were even discussed at a meeting of this Association six years ago [12]. Nevertheless, it seemed to me wise to begin the paper with a brief review of the elements of the theory, although I fear that the review may be too brief to be intelligible to those who are not already familiar with the material.

The consideration of a sequence of markets under conditions of uncertainty is not new in economics but does not seem to have received much attention from value theorists since the publication of Hicks's *Value and Capital* [11]. I would therefore have felt more comfortable presenting this paper in a session entitled, "Old Ideas in Pure Theory," but as far as I know, no such session has been organized for the current meetings.

I. Review of the Arrow-Debreu Model of a Complete Market for Present and Future Contingent Delivery

In this section I review the approach of Arrow [1] and Debreu [6] to incorporating uncertainty about the environment into a Walrasian model of competitive equilibrium. The basic idea is that commodities are to be distinguished, not only by their physical characteristics and by the location and dates of their availability and/or use, but also by the environmental event in which they are made available and/or used. For example, ice cream made available (at a particular location on a particular date) if the weather is hot may be considered to be a different commodity from the same kind of ice cream made available (at the same location and date) if the weather is cold. We are thus led to consider a list of "commodities" that is greatly expanded by comparison with the corresponding case of certainty about the environment. The standard arguments of the theory of competitive equilibrium, applied to an economy with this expanded list of commodities, then require that we envisage a "price" for each commodity in the list, or, more precisely, a set of price ratios specifying the rate of exchange between each pair of commodities.

Just what institutions could, or do, effect such

exchanges is a matter of interpretation that is, strictly speaking, outside the model. I shall present one straightforward interpretation, and then comment briefly on an alternative interpretation.

First, however, it will be useful to give a more precise account of the concepts of environment and event that I shall be employing. The description of the "physical world" is decomposed into three sets of variables: (1) decision variables, which are controlled (chosen) by economic agents; (2) environmental variables, which are not controlled by any economic agent; and (3) all other variables, which are completely determined (possibly jointly) by decisions and environmental variables. A state of the environment is a complete specification (history) of the environmental variables from the beginning to the end of the economic system in question. An event is a set of states; for example, the event "the weather is hot in New York on July 1, 1970" is the set of all possible histories of the environment in which the temperature in New York during the day of July 1, 1970, reaches a high of at least (say) 75°F. Granting that we cannot know the future with certainty, at any given date, there will be a family of elementary observable (knowable) events, which can be represented by a partition of the set of all possible states (histories) into a family of mutually exclusive subsets. It is natural to assume that the partitions corresponding to successive dates are successively finer, which represents the accumulation of information about the environment.

We shall imagine that a "market" is organized before the beginning of the physical history of the economic system. An elementary contract in this market will consist of the purchase (or sale) of some specified number of units of a specified commodity to be delivered at a specified location and date, if and only if a specified elementary event occurs. Payment for this purchase is to be made now (at the beginning), in "units of account," at a specified price quoted for that commodity-location-date-event combination. Delivery of the commodity in more than one elementary event is obtained by combining a suitable set of elementary contracts. For example, if delivery of one quart of ice cream (at a specified location and date) in hot weather costs \$1.50 (now) and delivery of one quart in non-hot weather costs \$1.10, then sure delivery of one quart (i.e., whatever be the weather) costs $\$1.50 + \$1.10 = \$2.60$.

There are two groups of economic agents in the economy: producers and consumers. A producer chooses a production plan, which determines his input and/or output of each commodity at each

date in each elementary event. (I shall henceforth suppress explicit reference to location, it being understood that the location is specified in the term commodity.) For a given set of prices, the present value of a production plan is the sum of the values of the inputs minus the sum of the values of the outputs. Each producer is characterized by a set of production plans that are (physically) feasible for him: his production possibility set.

A consumer chooses a consumption plan, which specifies his consumption of each commodity at each date in each elementary event. Each consumer is characterized by: (1) a set of consumption plans that are (physically, psychologically, etc.) feasible for him, his consumption possibility set; (2) preferences among the alternative plans that are feasible for him; (3) his endowment of physical resources, i.e., a specification of the quantity of each commodity, e.g., labor, at each date in each event with which he is exogenously endowed; and (4) his shares in producers' profits, i.e., a specification for each producer, of the fraction of the present value of that producer's production plan that will be credited to the consumer's account. (For any one producer, the sum of the consumers' shares is unity.) For given prices and given production plans of all the producers, the present net worth of a consumer is the total value of his resources plus the total value of his shares of the present values of producers' production plans.

An equilibrium of the economy is a set of prices, a set of production plans (one for each producer), and a set of consumption plans (one for each consumer), such that (a) each producer's plan has maximum present value in his production possibility set; (b) each consumer's plan maximizes his preferences within his consumption possibility set, subject to the additional (budget) constraint that the present cost of his consumption plan not exceed his present net worth; (c) for each commodity at each date in each elementary event, the total demand equals the total supply; i.e., the total planned consumption equals the sum of the total resource endowments and the total planned net output (where inputs are counted as negative outputs).

Notice that (1) producers and consumers are "price takers"; (2) for given prices there is no uncertainty about the present value of a production plan or of given resource endowments, nor about the present cost of a consumption plan; (3) therefore, for given prices and given producers' plans, there is no uncertainty about a given consumer's present net worth; (4) since a con-

sumption plan may specify that, for a given commodity at a given date, the quantity consumed is to vary according to the event that actually occurs, a consumer's preferences among plans will reflect not only his "tastes" but also his subjective beliefs about the likelihoods of different events and his attitude towards risk [16].

It follows that beliefs and attitudes towards risk play no role in the assumed behaviour of producers. On the other hand, beliefs and attitudes towards risk do play a role in the assumed behaviour of consumers, although for given prices and production plans each consumer knows his (single) budget constraint with certainty.

I shall call the model just described an "Arrow-Debreu" economy. One can demonstrate, under "standard conditions": (1) the existence of an equilibrium, (2) the Pareto optimality of an equilibrium, and (3) that, roughly speaking, every Pareto optimal choice of production and consumption plans is an equilibrium relative to some price system for some distribution of resource endowments and shares [6, Chaps. 5 and 6] [7].

In the above interpretation of the Arrow-Debreu economy, all accounts are settled before the history of the economy begins, and there is no incentive to revise plans, reopen the market or trade in shares. There is an alternative interpretation, which will be of interest in connection with the rest of this paper but which corresponds to exactly the same formal model. In this second interpretation, there is a single commodity at each date—let us call it "gold"—that is taken as a numeraire at that date. A "price system" has two parts: (1) for each date and each elementary event at that date, there is a price, to be paid in gold at date 1, for one unit of gold to be delivered at the specified date and event; (2) for each commodity, date, and event at that date, there is a price, to be paid in gold at that date and event, for one unit of the commodity to be delivered at that same date and event. The first part of the price system can be interpreted as "insurance premiums" and the second part as "spot prices" at the given date and event. The insurance interpretation is to be made with some reservation, however, since there is no real object being insured and no limit to the amount of insurance that an individual may take out against the occurrence of a given event. For this reason, the first part of the price system might be better interpreted as reflecting a combination of betting odds and interest rates.

Although the second part of the price system might be interpreted as spot prices it would be a mistake to think of the determination of the

equilibrium values of these prices as being deferred in real time to the dates to which they refer. The definition of equilibrium requires that the agents have the access to the complete system of prices when choosing their plans. In effect, this requires that at the beginning of time all agents have available a (common) forecast of the equilibrium spot prices that will prevail at every future date and event.

II. *Extension of the Arrow-Debreu Model to the Case in Which Different Agents Have Different Information*

In an Arrow-Debreu economy, at any one date each agent will have incomplete information about the state of the environment, but all the agents will have the same information. This last assumption is not tenable if we are to take good account of the effects of uncertainty in an economy. I shall now sketch how, by a simple reinterpretation of the concepts of production possibility set and consumption possibility set, we can extend the theory of the Arrow-Debreu economy to allow for differences in information among the economic agents.¹

For each date, the information that will be available to a given agent at that date may be characterized by a partition of the set of states of the environment. To be consistent with our previous terminology, we should assume that each such information partition must be at least as coarse as the partition that describes the elementary events at that date; i.e., each set in the information partition must contain a set in the elementary event partition for the same date.

For example, each set in the event partition at a given date might specify the high temperature at that date, whereas each set in a given agent's information partition might specify only whether this temperature was higher than 75°F. or not. Or the event partition at a given date might specify the temperature at each date during the past month, whereas the information partition might specify only the mean temperature over the past month.

An agent's information restricts his set of feasible plans in the following manner. Suppose that at a given date the agent knows only that the state of the environment lies in a specified set A (one of the sets in his information partition at that date), and suppose (as would be typical) that the set A contains several of the elementary events that are in principle observable at that date. Then any action that the agent takes at that

¹ This section is based upon [14, Sections 2-6].

date must necessarily be the same for all elementary events in the set A . In particular, if the agent is a consumer, then his consumption of any specified commodity at that date must be the same in all elementary events contained in the information set A ; if the agent is a producer, then his input or output of any specified commodity must be the same for all events in A . (I am assuming that consumers know what they consume and producers what they produce at any given date.)

Let us call the sequence of information partitions for a given agent his information structure and let us say that this structure is fixed if it is given independent of the actions of himself or any other agent. Furthermore, in the case of a fixed information structure, let us say that a given plan (consumption or production) is compatible with that structure if it satisfies the conditions described in the previous paragraph, at each date.

Suppose that the consumption and production possibility sets of the Arrow-Debreu economy are interpreted as characterizing, for each agent, those plans that would be feasible if he had "full information" (i.e., if his information partition at each date coincided with the elementary event partition at that date). The set of feasible plans for any agent with a fixed information structure can then be obtained by restricting him to those plans in the full information possibility set that are also compatible with his given information structure.

From this point on, all of the machinery of the Arrow-Debreu economy (with some minor technical modifications) can be brought to bear on the present model. In particular, we get a theory of existence and optimality of competitive equilibrium relative to fixed structures of information for the economic agents. I shall call this the "extended Arrow-Debreu economy."²

III. Choice of Information

There is no difficulty in principle in incorporating the choice of information structure into the model of the extended Arrow-Debreu economy. I doubt, however, that it is reasonable to assume that the technological conditions for the acquisition and use of information generally satisfy the hypotheses of the standard theorems on the existence and optimality of competitive equilibrium.

The acquisition and use of information about

the environment typically require the expenditure of goods and services; i.e., of commodities.

If one production plan requires more information for its implementation than another (i.e., requires a finer information partition at one or more dates), then the list of (commodity) inputs should reflect the increased inputs for information. In this manner a set of feasible production plans can reflect the possibility of choice among alternative information structures.

Unfortunately, the acquisition of information often involves a "set-up cost"; i.e. the resources needed to obtain the information may be independent of the scale of the production process in which the information is used. This set-up cost will introduce a nonconvexity in the production possibility set, and thus one of the standard conditions in the theory of the Arrow-Debreu economy will not be satisfied [14, Sec. 9].

There is another interesting class of cases in which an agent's information structure is not fixed, namely, cases in which the agent's information at one date may depend upon production or consumption decisions taken at previous dates, but all actions can be scaled down to any desired size. Unfortunately space limitations prevent me from discussing this class in the present paper.

IV. Critique of the Extended Arrow-Debreu Economy

If the Arrow-Debreu model is given a literal interpretation, then it clearly requires that the economic agents possess capabilities of imagination and calculation that exceed reality by many orders of magnitude. Related to this is the observation that the theory requires in principle a complete system of insurance and futures markets, which appears to be too complex, detailed, and refined to have practical significance. A further obstacle to the achievement of a complete insurance market is the phenomenon of "moral hazard" [2].

A second line of criticism is that the theory does not take account of at least three important institutional features of modern capitalist economies: money, the stock market, and active markets at every date.

These two lines of criticism have an important connection, which suggests how the Arrow-Debreu theory might be improved. If, as in the Arrow-Debreu model, each production plan has a sure unambiguous present value at the beginning of time, then consumers have no interest in trading in shares, and there is no point in a stock market. If all accounts can be settled at the beginning of time, then, there is no need for money

² This terminology is not in any way meant to imply that either Arrow or Debreu approve of this way of incorporating information into their model!

during the subsequent life of the economy; in any case, the standard motives for holding money do not apply.

On the other hand, once we recognize explicitly that there is a sequence of markets, one for each date, and no one of them complete (in the Arrow-Debreu sense), then certain phenomena and institutions not accounted for in the Arrow-Debreu model become reasonable. First, there is uncertainty about the prices that will hold in future markets, as well as uncertainty about the environment.

Second, producers do not have a clear-cut natural way of comparing net revenues at different dates and states. Stockholders have an incentive to establish a stock exchange, since it enables them to change the way their future revenues depend on the states of the environment. As an alternative to selling his shares in a particular enterprise, a stockholder may try to influence the management of the enterprise in order to make the production plan conform better to his own subjective probabilities and attitude towards risk.

Third, consumers will typically not be able to discount all of their "wealth" at the beginning of time, because (a) their shares of producers' future (uncertain) net revenues cannot be so discounted and (b) they cannot discount all of their future resource endowments. Consumers will be subject to a sequence of budget constraints, one for each date (rather than to a single budget constraint relating present cost of his consumption plan to present net worth, as in the Arrow-Debreu economy).

Fourth, economic agents may have an incentive to speculate on the prices in future markets, by storing goods, hedging, etc. Instead of storing goods, an agent may be interested in saving part of one date's income, in units of account, for use on a subsequent date, if there is an institution that makes this possible. There will thus be a demand for "money" in the form of demand deposits.

Fifth, agents will be interested in forecasting the prices in markets at future dates. These prices will be functions of both the state of the environment and the decisions of (in principle, all) economic agents up to the date in question.

V. *Equilibrium of Plans, Prices, and Price Expectations in a Sequence of Markets*

Consider now a sequence of markets at successive dates. Suppose that no market at any one date is complete in the Arrow-Debreu sense; i.e., at every date and for every commodity there will be some future dates and some events at those

future dates for which it will not be possible to make current contracts for future delivery contingent on those events. In such a model, several types of "equilibrium" concept suggest themselves. First, we may think of a sequence of "momentary" equilibria in which the current market is cleared at each date. The prices at which the current market is cleared at any one date will depend upon (among other things) the expectations that the agents hold concerning prices in future markets (to be distinguished from future prices on the current market!). We can represent a given agent's expectations in a precise manner as a function (schedule) that indicates what the prices will be at a given future date in each elementary event at that date. This includes, in particular, the representation of future prices as random variables, if we admit that the uncertainty of the agent about future events can be scaled in terms of subjective probabilities [16].

In the evolution of a sequence of momentary equilibria, each agent's expectations will be successively revised in the light of new information about the environment and about current prices. Therefore, the evolution of the economy will depend upon the rules or processes of expectation formation and revision used by the agents. In particular, there might be interesting conditions under which such a sequence of momentary equilibria would converge, in some sense, to a (stochastic) steady state. This steady state, e.g., stationary probability distribution of prices, would constitute a second concept of equilibrium.

I am not aware of any systematic general theory of markets under uncertainty, incorporating one or both of these two concepts of equilibrium, that has appeared since Hicks's *Value and Capital*, and I don't think that we can rest satisfied with Hicks's treatment in terms of "certainty equivalents" and "elasticities of expectation." The desirability of having a better theory and the importance of the role of expectations are well recognized, of course [3]. In the further development of such a theory, we shall no doubt have to face some of the difficult problems that have appeared in recent work on sequences of momentary equilibria under conditions of certainty [10] [17] [18].

A third concept of equilibrium emerges if we investigate the possibility of consistency among the expectations and plans of the various agents. I shall say that the agents have common expectations if they associate the same (future) prices to the same events. (Note that this does not necessarily imply that they agree on the joint probability distribution of future prices, since different

agents might well assign different subjective probabilities to the same event.) I shall say that the plans of the agents are consistent if for each commodity, each date, and each event at that date the planned supply of that commodity at that date in that event equals the planned demand and if a corresponding condition holds for the stock markets. An equilibrium of plans, prices, and price expectations is a set of prices on the current market, a set of common expectations for the future, and a consistent set of individual plans, one for each agent, such that, given the current prices and price expectations, each individual agent's plan is optimal for him, subject to an appropriate sequence of budget constraints.

Of the three concepts of optimality, the last is perhaps the closest in spirit to the Arrow-Debreu theory. How far do the conclusions of the Arrow-Debreu theory (existence and optimality of equilibrium) extend to this new situation? Concerning existence, for particular definitions of "individual optimality" and specifications of the agents' "budget constraints," one can prove the following theorem. Before stating the existence theorem I must define what I shall call a pseudo-equilibrium.

The definition of pseudo-equilibrium is obtained from the definition of equilibrium by replacing the requirement of consistency of plans by the condition that at each date and each event the difference between total saving and total investment (by consumers) is smaller at the pseudo-equilibrium prices than at any other prices.³

One can prove [15] that under assumptions about technology and consumer preferences similar to those used in the Arrow-Debreu theory: (1) there exists a pseudo-equilibrium; (2) if in a pseudo-equilibrium the current and future prices on the stock market are all strictly positive, then the pseudo-equilibrium is an equilibrium; (3) in the case of a pure exchange economy, there exists an equilibrium.

The crucial difference between this theorem and the corresponding one in the Arrow-Debreu theory seems to be due to the form taken by Walras' law, which in this model can be paraphrased by saying that saving must be at least equal to investment at each date in each event. This form derives from the replacement of a single budget constraint (in terms of present value) by a sequence of budget constraints, one for each date.

With regard to optimality, there is little that

³ This second condition will be automatically satisfied at an equilibrium. It should be noted that at each date the set of current prices is normalized; e.g., by taking the sum to be unity.

can be said at this time. The main difficulty in investigating this question seems to be in characterizing the set of states of the economy that are attainable, given the restrictions on the set of allowable contracts at each date.

VI. Unsolved Problems

I can only list here a few unsolved problems that I personally find interesting and promising for further research.

I have already mentioned the question of the optimality properties (if any) of an equilibrium of plans, prices, and price expectations. One possible approach is to consider more explicitly the information that the observation of prices provides for agents in the economy. One might hope to show that an equilibrium is an optimum relative to the set of states of the economy that could be attained with just the same information that is provided by the equilibrium prices (in addition, of course, to the information structures originally available to the individual agents). Notice that since the equilibrium price expectations are self-fulfilling, the observation of the prices in any current market provides information about the true state of the environment (i.e., the specification of the values of particular prices defines an "event" in the set of possible states of the environment). An approach of this kind was tried in a two-period model [13], which was further complicated, however, by allowing agents to make contracts for future delivery contingent on the values taken on by future prices. (An example of such a contract would be a wage contract with a cost-of-living escalation clause.) It was shown that in this model, if the introduction of such contracts enabled all the agents to discount future receipts and costs back to the initial date (i.e., if all uncertainty about the environment could be reflected in some corresponding uncertainty about future prices), then an equilibrium would be an optimum in the above sense. Unfortunately, the existence of an equilibrium in such a model was not demonstrated, and indeed there might be important economic phenomena that would rule out the existence of equilibrium in such a model [13].

I have also already mentioned the unsatisfactory state of the theory of the evolution of momentary equilibria in a sequence of markets and the question of possible convergence of momentary equilibria to a (stochastic) steady state.

In all of these (potential) theories of a sequence of markets we shall need a more detailed theory of the firm than that used in the Arrow-Debreu model. Simple profit maximization is not well defined if future profits are uncertain and

cannot fully be discounted back to the present. The model of Section V essentially begged this question by assuming that each producer maximizes a utility function whose arguments are his future net revenues in different events. Such an assumption fails to relate the behavior of the firm to the preferences of the stockholders or potential stockholders. (It is rather an expression of the "divorce of ownership from management"!) An alternative candidate that has been discussed is the assumption that at each date a producer maximizes the current stock market value of his firm. (Note that in the Arrow-Debreu model, profit maximization is equivalent to maximization of the value of the stock.) However, except in the context of a special example considered by Diamond [8], I have not seen a formulation of this hypothesis that enables the producer to act as a price-taker; i.e., that does not imply that the producer is able to calculate the effect of his actions on the equilibrium prices.

We shall also want to incorporate into our theories the process of entry and exit of firms. In particular, the results described in Section V on the relationship between equilibrium and pseudo-equilibrium suggest that the possibility of exit may be important in assuring the existence of such an equilibrium.

Finally, I mention the old problem of incorporating a theory of money and credit in a Walrasian model of general equilibrium [9]. In a sense, the model of Section V allows "secured" loans that are backed either by physical collateral or by contracts for future delivery of commodities. The theory also provides a framework for explaining the holding of "commodity money." The model does not, however, describe any institutions for carrying over "units of account" from one date to the next; the introduction of such institutions seems a natural next step and one for which the model seems to me to be well suited.

REFERENCES

1. K. J. ARROW, "Le Rôle de Valeurs Boursières pour la Répartition la Meilleure des Risques," *Econométrie*, 1953 (Centre National de la Recherche Scientifique), pp. 41-48; or see the translation, "The Role of Securities in the Optimal Allocation of Risk Bearing," *Rev. of Econ. Studies*, 1964, pp. 91-96.
2. ———, *Aspects of the Theory of Risk-Bearing* (Helsinki: Yrjö Jahnsson Lecture Series, 1965).
3. M. J. BOWMAN, ed., *Expectations, Uncertainty, and Business Behavior* (S.S.R.C., 1958).
4. D. CASS and J. E. STIGLITZ, "The Implication of Alternative Saving and Expectations Hypotheses for Choices of Technique and Patterns of Growth," *J.P.E.*, July-Aug., 1969, pp. 586-627.
5. G. DEBREU, "Une Economie de l'Incertain" (Paris: 1953, Electricité de France) (mimeographed).
6. ———, *Theory of Value* (Wiley, 1959).
7. ———, "New Concepts and Techniques for Equilibrium Analysis," *Int. Econ. Rev.*, 1962, pp. 257-73.
8. P. DIAMOND, "The Role of a Stock Market in a General Equilibrium Model with Technological Uncertainty," *A.E.R.*, Sept., 1967, pp. 759-76.
9. F. H. HAHN, "On Some Problems of Proving the Existence of an Equilibrium in a Monetary Economy," in Hahn and Brechling, eds., *The Theory of Interest Rates* (London: Macmillan, 1965).
10. ———, "Equilibrium Dynamics with Heterogeneous Capital Goods," *Q.J.E.*, Nov., 1966.
11. J. R. HICKS, *Value and Capital* (Clarendon Press, 1939).
12. J. HIRSELEIFER, "Efficient Allocation of Capital in an Uncertain World," *A.E.R.*, May, 1964, pp. 77-85.
13. R. RADNER, "Equilibre des Marchés à Terme et au Comptant en Cas d'Incertainité," *Cahiers d'Econométrie* (Paris: Centre National de la Recherche Scientifique, 1967).
14. ———, "Competitive Equilibrium under Uncertainty," *Econometrica*, Jan., 1968, pp. 31-58.
15. ———, "Equilibrium of Plans, Prices, and Price-Expectations in a Sequence of Markets" (1969, unpublished).
16. L. J. SAVAGE, *The Foundations of Statistics* (Wiley, 1954).
17. K. SHELL and J. E. STIGLITZ, "The Allocation of Investment in a Dynamic Economy," *Q.J.E.*, Nov., 1967, pp. 592-609.
18. H. UZAWA, "Market Allocation and Optimum Growth," *Australian Economic Papers*, June, 1968, pp. 17-27.

DISCUSSION

RUTH P. MACK: The new ideas we have been hearing about follow, as I see it, the basic strategy of neo-classical work in the following sense: Classical economic theory started with a set of assumptions and built thereon a theory about how the production and distribution of goods should be organized in order to optimize economic well-being. The assumptions to which I refer postulate (1) economic man, (2) a highly competitive market structure, (3) rich and sure information, (4) scarce resources. Neoclassical work among other things softened assumption 2; competitive markets were modified to admit monopolistic elements. The received manipulative apparatus of theoretical economics was employed, and honed and developed thereby, to work through the implications of the relaxed assumptions. The implications were focused in the traditional way on how equilibrium tendencies would assert themselves and the extent to which results would prove optimal.

The new ideas discussed in this session derive from further modification of the classic assumption about the economic world. In line with other work of recent years, assumption 3 has been modified to convert sure information into probabilistic, partial, and even ambiguous information—the foundation on which Radner's paper builds. In game theory, assumption 2 has been further rewritten to modify the structure of competitive markets to include coalitions and assumption 1 is also altered—economic man becomes a collective of economic men, as Shapley's paper illustrates. Both authors find that they can under some conditions win through to rules that specify equilibrium and even optimal solutions. Under other conditions the road is blocked by partial (and therefore unshared) information in Radner's analysis and by externalities which are diseconomies in Shapley's.

The work then is in the grand tradition of crafting, under the building codes of main-line economic theory, economic construction jobs grounded on a highly selective modification of the classic assumptions.

But though Roy Radner revises assumptions about information, he leaves the others in their ancient garbs. Yet it is material to his analysis that natural man is in essence, both as consumer and producer, a fundamentally different animal from economic man: (a) his perception is not catholic but selective and preorganized; (b) he is characterized by a developmental and indeed a partially elective level of aspiration; (c) his value schemes are likewise developmental; (d) economic values are only one part of the value system pertinent to economic behavior; (e) the decision agent is typically not a man but a group of men, a collective. Modifications (a) to (c) provide the foothold of learning. They would, for example, challenge whether "information structure" of which Dr. Radner speaks can in fact be "fixed" in the sense that "it is given independently of the actions of himself or any other agent."

Yet perhaps the most serious implications of revision

of the classical assumptions lie in the way in which relaxations in one assumption interact with relaxations in others. For example, a very large firm makes an important investment decision (revised assumption 2) under sparse and probabilistic information (revised 3) based largely on the outcome of a power struggle among board members (revised 1_a and 1_e), which then influences a flock of investments by other firms in a rolling cumulating sequence. A harder look at assumption 4, combined with revisions in 1 and 2, indicates that resources can be so greatly enhanced, at so small and erratic an opportunity cost, as to suggest they should not be viewed even in theory as "scarce." Interactions of these sorts introduce a fundamental unknowability into trends and even types of change in the long run.

Space forbids more than these few cryptic remarks. However, note that I have been speaking of relaxations in assumptions the net result of which is not simply to smear the outlines of an equilibrium path (make it a swathe rather than a line) but rather to bias results (bend or even angle the path). I conclude that general long-term equilibrium, much less optimal tendencies, cannot be identified in the polluted atmosphere of realistic (however grossly simplified) assumptions.

The dilemma is clear. If theoretical analysis accepts the challenge of broadly realistic assumptions, the search for grand equilibrium optimal solutions will usually need to be abandoned. If it rejects the challenge, theory will regress into an ever deepening scholasticism. Possibly some revolutionary new paradigm will resolve the conflict. But in the meantime selective approaches may prove useful—approaches that utilize the rigorous methods, tools, and concepts of economic analysis to explore real-world problems.

Selective approaches deliberately match the method to the realities of the problem. An important aspect of the matching resorts, in effect, to a preference surface that trades generality for realism. In some connections generality can be purchased at very little sacrifice of realism. In others quite the reverse applies. It is essential to identify these situations. This is of the essence.

The most obvious identifiable situations in which the cost of realism is low are those for which the relevant realistic assumptions are close to those of the classic model and short-run tendencies are of interest. Problems concerning scrap markets, open exchange markets, and some aspects of international trade are among the obvious candidates.

Selection depends also on the effectiveness with which progress in theoretical analysis can achieve economies in the cost of generality in terms of realism. The new ideas exemplified in this session have hinted at ways in which such economies might be achieved.

One approach consists of the design of new analytic tools and concepts. Game theory seems to be one such innovation (and, I might add, information theory another). The notion of the "core," particularly,

promises to be a remarkably flexible yet disciplined implement. The Shapley-Shubik paper recognizes that, even when the existence of a core cannot be proved, the notions of the core provide "an investigative tool, for case-by-case application, that can give important insights into the prospective stability of cooperative internalization of external diseconomies."

A further suggestion based on a possible analogy to the way in which economists have learned to use the brilliant notion of a schedule: Demand and supply schedules adapt themselves to thinking about price histories in three steps: (1) the shape of the schedules, (2) the institutionally determined flexibility of prices, (3) change in the schedules. The beauty of the construct is displayed by its high usefulness even when most of the actual history of prices is due to change in schedules.

Am I forcing the point on Roy Radner by interpreting his suggestions about limited sorts of equilibrium analyses in these terms? His "steady state" equilibrium falls in step 1. Here assumptions may appropriately be purified of their dynamic real-world attributes. The latter are added as a related step 3, in which one asks how the steady-state path would in fact shift.

His other suggestions—an equilibrium of plans, prices, and price expectations—seem to imply concentration on step 2—the nature of equilibrium convergence tendencies. However, here very particular care is required—the avoidance of internal inconsistency of assumptions. I am troubled enough about such inconsistency in the Arrow-Debreu use of period analysis in which they assume that plans of producers are fixed without response to the relation between consumers' probabilistic expectations and unfolding reality. But the difficulty is underscored when Roy Radner takes the daring (and valuable) step of admitting differences in information among consumers. For the very possibility of differing information implies that the whole process of the formation of expectations about future events and participants' actions becomes vastly more complicated and interactive and applicable necessarily to producers as well as consumers. Period analysis itself tends to harbor such internal inconsistencies when it forces into one "period" the relationship among processes for which natural periods are actually very different: change in production of finished goods, production planning (fixed to rule out inventory change), consumer planning, the formation and revision of expectations, adjustment of prices to demand-supply relationships, etc.

In general, I am suggesting that if the aspiration toward grand economy-wide equilibrium optimality could be abandoned and internal inconsistency in assumptions minimized, it seems clear that the theorist's kit of tools will in the hands of the fine craftsman to whom we have just listened turn out, on a selective basis, rigorous and urgently needed realistic analysis of economic processes.

KENNETH J. ARROW: We are indeed fortunate to have had three papers which combine so well depth of content with clarity of exposition. The authors have

between them dealt with three areas in general equilibrium theory which have shown the most rapid development.

Scarf's paper is a continuation of the work on existence theorems for competitive equilibrium but represents a vast step forward. Walras was already concerned with the difference between demonstrating the existence of equilibrium and finding a method of achieving it. Walras and virtually all subsequent writers in one form or another sought an approach by guessing a price vector and then using supply-demand information to get better and better approximations. But Scarf showed some years ago that this approach could not be valid in general; that is, he exhibited an economic system in which the unique equilibrium was unstable under a dynamics of adjustment of each price to excess demand on that market. His present approach is very different; most strikingly, we find that at each stage we have not a single price vector but a set of price vectors (which are, however, close to each other). Unfortunately there seems at the moment to be no obvious economic interpretation of the successive steps.

However, I do want to point out that Scarf's algorithm should shortly find empirical application. General equilibrium theory has justly been attacked for having found so far little use in direct application to empirical situations. But this is not a matter of principle but of difficulty in making the application. The qualitative theorems with which we are accustomed in partial equilibrium analysis turn out for the most part to be false in general equilibrium theory; hence, its predictions must take quantitative form; i.e., inferences from specific production and demand functions to equilibrium prices and quantities. For this inference, an algorithm is needed and is now supplied. The most obvious field of application is the equilibrium of international trade. Ever since the Leontief paradox was propounded, the usual explanations have run in terms of a three or more factor version of the Heckscher-Ohlin model. But it has been impossible to confront the extended model with the facts in a meaningful way. Instead, some simple correlations between, e.g., research-intensity and percentage of produce exported are calculated. But exports will in fact depend on the factor endowments of other countries as well as of the exporting country, and such relations can only be computed from a full general equilibrium model. It is here that Scarf's algorithm should be most useful.

The concept of the core is one of the most exciting developments in equilibrium theory in years, and its full implications have yet to be understood. Shapley and Shubik have given, in addition to a most useful statement of the concept, an extension to the difficult problem of externalities. I would like to raise one question, the answer to which I am not at all sure of. They suggest that there may be a sharp distinction between positive and negative externalities. They argue that if the former are internalized, by being listed as explicit commodities, the core will exist, a conclusion which gives some support to the idea of Coase

and others that unrestricted bargaining will lead to Pareto efficient outcomes even in the presence of externalities. But they give examples to show that in the presence of negative externalities, the core may not exist. The question is, whether the difference may have to do with convexity assumptions rather than directly with the sign of the externality. My colleague, David Starrett, has pointed out that, with respect to production, a negative externality is incompatible with concavity of the production function.

If, for example, output of a firm is negatively related to pollution of water and air by others, concavity implies increasing effect of pollution as its magnitude increases, and therefore the firm would necessarily have zero output at a finite pollution level and negative outputs beyond. But since the firm can always shut down and have zero output at any pollution level, it is impossible to have concavity throughout. Also, Debreu has shown that admitting negative prices does not by itself destroy the possibility of competitive equilibrium. Hence, if all externalities were listed as commodities and if convexity assumptions prevailed in the enlarged commodity space, there would be a competitive equilibrium which should be an element of the core.

Radner's paper covers the field of equilibrium under uncertainty so thoroughly that as a discussant I am reduced to some very small observations: (1) Radner gives lack of information and moral hazard as two distinct reasons for the failure of some markets for contingent claims to exist. But in fact the latter is a special case of the former; if an insurance company could distinguish whether a fire was due to arson or not, it could pay in the latter case but not in the former. Thus moral hazard arises only because the insurance company cannot distinguish between two states of nature. (2) Under uncertainty, the securities of different firms are no longer perfect substitutes. Radner suggests pessimistically that we can no longer speak of perfect competition. But in the spirit of the Modigliani-Miller model, we can imagine large classes of firms which have the same risks, so that the possibility of perfect competition is restored, while still exposing firms to risk and not all firms to the same risks.

TROUT RADER: Shapley and Shubik consider the core as a concept for social stability. It is argued here that the core is of great interest mainly to welfare economics, not to positive social analysis.

The main points made by Shapley and Shubik are: (1) Economies with externalities due to inappropriabilities may have cores (or Edgeworth optimal allocations); namely, the competitive equilibrium with externalities appropriated. (2) There is asymmetry between the core of "negative" and "positive" externality in that only in the latter case will there necessarily be a core. An additional point not observed by Shapley and Shubik is as follows: whether there are negative or positive externalities, there will be a competitive equilibrium. Goods of negative value will have negative prices and the core will exist if the "game" forbids all transfers which are not mutually

voluntary. For purposes of discussion, call this competitive equilibrium the "ideal state."

According to a conjecture of Edgeworth¹ which Scarf² proved, the ideal state and the core correspond for economies where it is physically possible for groups to trade only among themselves. (In this way the old distribution problem of welfare economics was solved—to a limited extent.) Therefore, the asymmetry observed by Shapley and Shubik holds only if nonvoluntary transfers are allowed. Only then it is possible to have no core. However, there are still Edgeworth optimal points in the sense of welfare economics, since if groups could be made independent, they could not do better than the ideal state. For instance, in the garbage game, the ideal state is to leave garbage where it is. If a group need not worry about the dumping of others, it could do no better than the ideal state. Without policing, the ideal state is not in the core and, it is said, the ideal state is not "socially stable." However, the symmetry of the game makes the latter conclusion highly suspect.

If a solution is not in the core, there will be a temptation to disturb it. The aggrieved group has obvious gains of disturbance. However, these gains are not to be expected with certainty whenever there is ambiguity as to which coalition will break off. For example, in the garbage game, no one can count on being in a majority coalition with probability one, whereupon with some positive probability he will be worse off than in the ideal state.

In a game with identical players, all coalitions with the same number of participants would be equally probable. Each player appears in the same number of such coalitions and has the same return to nonagreement. By the categorical imperative, each party will be morally constrained from submitting to divisionary tactics. In the case of risk aversion, the returns can be shown to be less than the ideal state. Otherwise, the return to nonagreement is no less. (Whatever the probabilities of outcome symmetry applies to show that the expected return is equal to b .)

The lake game also can be solved by symmetry considerations. By the same categorical imperative argument, the ideal state gives equal payoffs and it is "morally" viable. However, the desymmetrized model is more complicated.

If firms are charged for pollution to others, they will install a control apparatus. There is an ideal state as before, but it is not clear that it is viable. For example, consider two firms who pollute enough together so that in combination it pays to install a purification system. Assume that no one of the pair is important enough to purify in combination with a third, small firm, i . Hence, i can always count virtually a free ride in case of nonagreement. He must then be paid a fee greater than in the ideal state. Far from favoring the large firm, application of the categorical

¹F. Y. Edgeworth, *Mathematical Psychics* (Kegan Paul).

²Herbert Scarf, "An Analysis of Markets with a Large Number of Participants," *Recent Advances in Game Theory* (Princeton Univ. Press, 1962).

imperative favors the smaller one. (This is a familiar problem for industries where several large firms enforce monopoly power while the smaller firms do not constrain their output and still reap the benefits of monopoly.)

In contrast, it is possible for the greater polluters to follow the "immoral" policy of an optimal threat. They can threaten to pollute if there is no agreement among the three and can gain a larger share for themselves than the ideal state. If nonagreement were decided upon once and for all, such a threat would not be credible. However, should the larger firms act in nonagreement so as to force contract at their terms, the threat could be carried out for a short period of time (although it may not be very convincing).

To summarize, in the lake pollution game, there is a core, an ideal state, an agreement with credible threats, and an agreement with optimal threats. As

Shapley and Shubik mention, the core tells us nothing about viability since it is so large. The others have radically different outcomes and it is hard to see which is socially viable. It seems safe to infer only that the ideal state is not viable. The other two (or combinations between them) would depend upon the morality and credulity of the participants.

The same remarks apply to the smelting game except that generally there is no core! Nevertheless, it can be shown that there are outcomes which are socially viable for credible threats.³

It is hard to escape the conclusion that the core is not a proper solution concept for games. After some rules of property ownership are specified, it is a concept for welfare economics.

³ Trout Rader, *Theory of General Economic Equilibrium*, Chap. 3 (forthcoming, Academic Press).

THE ECONOMY OF CITIES*

DISCUSSION

KARL A. FOX: Most of my comments will relate to the paper by Jane Jacobs. Mrs. Jacobs is to be commended for her concern about urban problems and also for attempting to think about these problems in a systematic way. However, she gives us no definition of a city nor does she seem to recognize that a definition is needed. Lacking insight into the economic structure of the basic unit of her system, she also fails to provide us with useful criteria for policy intervention.

I cannot operate with Mrs. Jacobs' categories. We need a much better framework for conceptualizing urban economic structure than she has given us. The key to this better framework lies in central place theory, a field actively pursued by quantitative geographers but strangely neglected by most economists. I will begin by presenting my view of the structure of the United States economy as a hierarchy of central places.

Quantitative geographers recognize about eight levels of central places in the United States: hamlets, villages, towns, small cities (e.g., county seats), regional capitals, regional metropolitan centers, national metropolitan centers, and the national (economic) capital.

Hamlets and villages are relics of the horse-and-buggy era. Towns (perhaps 1,000 to 5,000 people), small cities (perhaps 5,000 to 25,000 people), and regional capitals (usually with 25,000 to 250,000 people) form a three-level hierarchy which (together with the farm and nonfarm village, hamlet, and open-country population) forms the basic ecological pattern of our nonmetropolitan areas in the automobile age.

The regional capital is the central city of a home-to-work commuting field of approximately an hour's radius. Since 1961, I have been calling such a commuting field a "functional economic area" or FEA. Such an area is a relatively self-contained labor market in the short run. It is also relatively self-contained in terms of other characteristics and systems such as department store trade areas, radio and TV broadcasting areas, newspaper advertising areas, community colleges, area vocational-technical schools, and medical and legal services.

A functional economic area may be regarded as a low-density city or rural-urban synthesis covering up to 5,000 square miles, equivalent to several midwestern counties. The regional capital contains the central business district of the FEA as a whole, as well as the largest and most diversified array of employment op-

portunities. Virtually all residents of the FEA are aware of the central city and make some use of it. In the Midwest, the total populations of FEA's will usually range from 100,000 to 500,000, even though their regional capitals as such contain only 25,000 to 250,000 people.

Ordinarily, few workers residing in one FEA commute to jobs located in other FEA's. Perhaps 60 percent of all jobs in an FEA are oriented to the needs of its own residents: retailing, public services, personal and professional services, and the like. The remaining 40 percent or so of jobs in an FEA are in activities which are primarily export-oriented from the standpoint of that FEA.

Two FEA's with the same population size and average income level will have almost identical arrays of resident-oriented jobs; the resident-oriented components of the two labor forces should be almost as interchangeable as the crews of two sister ships. National chains of department stores and national advertising of consumer goods both reflect and reinforce this basic similarity and interchangeability. The two FEA's may differ substantially in their arrays of jobs which are primarily export-oriented. Agriculture, most mining and manufacturing, large federal installations, state capitols and state universities (among other activities) are primarily export-oriented from the standpoint of the FEA in which they are located.

The complete distributions of wage and salary incomes, occupational skills, educational attainments, and potential community leadership in two such FEA's result from their similar arrays of resident-oriented jobs and their more or less dissimilar arrays of export-oriented jobs.

The functional economic area is the basic building block of the national system of cities in the United States. The resident-oriented activities and individual households in each FEA can be regarded as forming a sort of "macro-household"; the export-oriented activities form an intercity trading and input-output system which links the several hundred FEA's into a truly national economy.

The three-stage central place hierarchy of a nonmetropolitan FEA (town, small city, and regional capital) has its analog in the three-stage hierarchy of shopping centers found in the larger metropolitan areas: neighborhood shopping centers (serving from 15,000 to 30,000 people), district shopping centers (serving perhaps 60,000 to 120,000 people), and regional shopping plazas (serving as many as 500,000 people). A regional shopping plaza usually contains a major branch of a large downtown department store; few residents of the trade area of the regional shopping plaza make much use of the shopping facilities in the central business district of the metropolis as a whole.

Therefore, as consumers, most residents of the major metropolitan areas are clustered into functional

* The papers by Jane Jacobs, "Strategies for Helping Cities," Barbara R. Bergmann, "The Urban Economy and the Urban Crisis," and Anthony Downs, "Housing the Urban Poor: The Economics of Various Strategies," were printed in the Sept., 1969, issue of *A.E.R.* but presented at this session.

economic areas in the 200,000 to 500,000 population range. The metropolitan area as a whole may be viewed as equivalent to the close packing of several or many FEA's with their patterns of commuting to export-oriented jobs scrambled and distorted by rapid transit systems and freeways.

Thus, we arrive at a modular conception of the structure of the United States economy in terms of approximately 500 FEA's, each containing either a regional capital as such or a regional shopping plaza and other facilities normally found at the regional capital level. These FEA's and metropolitan subareas are the strategic areas for "community control" of resident-oriented public services.

The drives for "new towns," "experimental cities," and "rural-urban population balance" should be channeled into the planning and development of selected nonmetropolitan FEA's to bring their populations up from 100,000 or 200,000 currently to (say) 300,000 or more a decade or two hence. Redevelopment of metropolitan areas should reinforce the geographic integrity and sense of community of relatively self-contained subareas of not more than 500,000 people.

Both types of communities have tremendous untapped potentials for meeting human needs. Mrs. Jacobs' emphasis on "import-replacing" and "export-generating" activities implies a preoccupation with national and regional metropolitan centers each of which is a conglomerate of several or many FEA's. Indicators suggesting wholesome economic growth for such a conglomerate may disguise major problems which can only be dealt with at the subarea (FEA) level.

DAVID S. MUNDEL: The city is the locale of many of America's most pressing domestic problems for a multitude of reasons. Economics and economists can, and probably will, play an increasingly important role in our society's efforts to improve its urban condition. This role is one of increasing our understanding of the mechanisms and processes of urban behaviors as well as developing policies with which the society can fruitfully change these behaviors. The three papers presented here provide us with a view of the shape this role may take. They provide us with a diversity of descriptions of the operations of the "economy of cities" and of the limited utility of our current efforts to solve the "urban crisis." The papers do not provide us with the much-needed view of the underlying structure and basic causes of our urban problems. Without such a structure, it is unlikely that economists' solutions to the "urban crisis" will be any more productive than those developed by the research efforts which have preceded economics into the urban arena. The papers, however, do lead implicitly in the direction of a more structured view of the problems of the urban economy. This structure is their most important joint product.

The city is the locus of problems, because the American society has failed, in large part, to act effectively in two major areas of social policy. These areas are basic subjects of inquiry in any introductory public finance or welfare economics text, but when economists move to the urban frontier of their disci-

pline, they seem to be virtually ignored or forgotten. The first of these is the problem of income redistribution. The second is the activist regulation of men's interference with one another—externalities. These papers allude to these concerns in a variety of ways. They point to the disfunctional impacts of existing attempts to deal with them in an urban society. Today's urban crises are the result of our failure to effectively redistribute income and regulate externalities. Our continuing failure to understand this will result in our inability to solve these basic social problems.

The redistribution (or welfare) policies of America have actively encouraged the urbanization and segregation of low-income individuals and inhibited the economic vitality and viability of the urban areas within which they have relocated. The current forms of public housing and welfare systems demand large bureaucratic organizations for their management and these can only be supported in large urban centers. Thus, in order to receive substantial subsidies, low-income individuals must move to urban areas. The two national strategies that Jane Jacobs identifies may provide funds to an urban economy (often at the expense of another), but it is unlikely that they provide much support to that part of the urban economy—the poor—which is the real locale of the nation's urban crisis. The employment and other income effects of programs which support high-technology industries, highly bureaucratized service sectors, and physical renewal efforts are of minimal importance to the urban poor especially if they are systematically restricted from the labor markets which supply these sectors of the economy. They are therefore of minimal help to cities if the poor are where the problems are. Not only do these programs have minimal effects on the improvement of the condition of the urban poor, but they may also actually inhibit, in the long run, the economic vitality of the city to such an extent that their condition is worsened.

A welfare or redistribution system which provides income-in-kind and bureaucratizes the cash payment system may so substantially inhibit the economic development of a city that its residents can no longer improve their condition. The Jacobs-described processes of "adding new work to old" and "import replacement" are unlikely to take place in an economy in which decisions are increasingly the province of bureaucracies rather than of individuals. An estimate of this likelihood (or unlikelihood) is provided by a brief digression to the story of Mrs. Ida Rosenthal as described in Mrs. Jacobs' newest book, *The Economy of Cities*.

Mrs. Rosenthal designed and marketed a new brassiere because she was disappointed with the shape of the dresses purchased by the customers of her small New York dress shop. If Mrs. Rosenthal's customers had had to receive the approval of the Bronx Reclothing Authority (BRA) prior to their purchase of one of her dress and brassiere combination garments, it is unlikely that Mrs. Rosenthal's Maidenform Brassiere Company and the resulting economic spillovers (or the economic creativity as Mrs. Jacobs de-

scribed it) would have ever come into being. The same lack of incentives for change, development, and performance are found throughout the systems which provide housing, food, medical care, and schooling to the urban poor. In his paper on "Housing the Urban Poor: . . .," Anthony Downs states that "clearly, the bigger the total direct public subsidy put into housing, the more new units will be built." Regrettably, the current systems of housing subsidization make the validity of Downs's statement somewhat less than clear.

Thus, America's current redistribution policies are less than satisfactory for two reasons: (1) They fail to redistribute a large portion of their funds to the poor. (2) They set up incentives and decision-making processes which inhibit the economic growth and development of the cities in which the poor are located. These failures tend to exacerbate the problems of the "economies of cities" because one of the basic problems is urban poverty.

The failure of society to actively regulate the interference of men with one another is the second basic cause of American urban problems. Economics and the American society grew and developed at a time when these interferences were of minimal importance. But the current urban problems of pollution, crime, congestion, and discriminatory market restrictions are clearly anything but the perfectly competitive, no externalities worlds of our upbringings. It is the role of economists to locate and measure the externalities or conditions of well-being of the urban environment and to aid in the search for measures to compensate for and regulate their influences.

Barbara Bergmann's attempt to develop an accounting system which provides a systematic view of the connections between the urban economy and the social and environmental problems which confront the city is an important step in providing a structure within which our search for knowledge of "urban externalities" can be conducted. The "problem solving" information system which Mrs. Jacobs describes seems likely to take the form of Mrs. Bergmann's input-output accounting system and to be a useful instrument of the urban policy-maker. It is time for economists to analyze and work on the design of a process for getting somewhere (i.e., redesigning city policy-making process to make it a learning process), in addition to applying their efforts to their characteristic work of saying where our society ought to be. Downs's paper also alludes to problems of externalities. Downs reviews the restrictive effects of many of the public efforts to improve the housing of the urban poor, but he almost completely neglects to define the social interests in private housing decisions. Building codes restrict the housing options of the urban rich and poor alike, but do the restrictions create any social benefits? Who receives these benefits and at what cost to whom? Restrictive zoning enables suburban communities to avoid the external diseconomies of observing the residences of the poor, but at what cost to the poor and to their cities?

In brief, the problems of the American city are in large part the result of society's inept and often inap-

propriate solutions to the problems of income distribution and interaction effects. These are the basic issues of welfare economics. It is not surprising that when economists are concerned with the well-being of part of our society, then this area of the discipline should provide a productive structure for their concern and efforts. It is surprising that so many of us have continually looked elsewhere.

SAM BASS WARNER, JR.: Since my work as a historian has largely dealt with the issues raised by Mr. Downs and since I find very interesting and important policy considerations in his paper, I will confine myself to a historian's view of the metropolitan location of the poor, and the trickle-down housing market.

I do want, first, however, to call your attention to the extreme pessimism and caution of the papers at this session. This pessimism, marked by calls for data gathering, decentralization, and reduction in housing quality standards, I understand to be part of the general gloom which has settled over all urban reform sessions in the past few years. The Vietnam war, the unceasing military budget, and repeated cuts in civilian programs, carried forth in the face of mounting urban needs, have knocked the starch out of what has generally been one of the most energetic, even utopian, fields of discussion, planning, and reform. I find it hard, as others on this panel obviously do, to specify an urban future while the wealth, good will, and dreams of the nation are swallowed up in imperialist war and endlessly ramifying military and space projects.

From the historical point of view, Mr. Downs's placement of the poor in former middle-income housing at the core of a city applies only to the years subsequent to the middle of the nineteenth century. Prior to that time the poor could and did locate at the fringe and by walking one to three miles had a wide choice of jobs. The poor became concentrated because of the central location of old housing and because of their need for a central location to get satisfactory job access. Since nineteenth century street railway transportation took a radial form and since until the 1920's the greatest concentration of urban economic activity lay at the center, the poor became progressively confined to a core location. They did so despite very high land costs which forced them to pay high rents and endure extreme crowding of their residential structures. Only in industrial satellites, like Gary, Indiana, or Newark, New Jersey, did the poor also settle in large numbers. Here their hope was that the many mills and large size of satellite plants would offer job opportunity. It might be well to remember in this connection that before the automobile became a tool of the lower class, firms seeking fringe locations—steel mills, cement plants, ship-builders, refineries, etc.—often had to either provide housing for their workers or subsidize the extension of public transportation.

Now with the growth of the multicentered, highly differentiated metropolitan regions in the twentieth century there should have been a diffusion of the location of the poor out through the aging suburbs of

the metropolis as they sought access to jobs. Rising racial prejudice in the twentieth century has prevented this movement. Racial segregation was more intense in 1930 than in 1900, more intense still in 1960 than 1930. The result of this prejudice, both in closing Negroes out of jobs and in preventing new residential location, has been the current urban novelty of pinning the poor to the old core cities, the concentration of low paying jobs at the core, and the sick social cycle of low paying jobs, high unemployment, bad housing, and bad municipal services with which we are all familiar.

I conclude from this historical review that any transportation policy and any employment policy which will increase the job access of the poor throughout the metropolitan region will help to restore this class at least to their former nineteenth-century opportunities, although this is indeed a sad goal for a progressive economy and nation.

Mr. Downs's model of the trickle down of middle-income housing units to the poor also ignores an important issue of what we might call the unavailability of residential social overhead capital. Core slums and the mass of cheap new housing developments on the urban fringe share a common disability—both lack, and always have lacked, decent public facilities. Harlem was a special case: it once was a moderate to high standard area and was designed and served as such. It is an example of Mr. Downs's system of the urban poor's inheriting formerly adequate areas which they use up and the city neglects. But succession to decent middle-class districts is a special case in urban housing. The more frequent case is that of a cheap working class or lower-middle-class development which was not adequate at the very moment it was built. The new tenements built in 1895 in the Lower East Side of

New York, the mill housing built in Gary, Indiana, in 1915, the tiny single house bungalows of fringe automobile plant districts in Detroit, or the latest cheap subdivision in St. Louis lack important utilities like public transportation, sewers, and water, lack good land plans that will offer all houses dry sites, and proper orientation to the sun, storms and wind, lack safety from traffic, lack good schools, and lack decent residential services for recreation and shopping. There was typhoid and tuberculosis on the Lower East Side and only the rough care of the old municipal dispensaries to help. Today the medical care of the working class and old people who live in the urban fringe is a scandal.

What I'm trying to say is that in the past, as today, there was precious little residential social overhead capital to trickle down to subsequent occupants. Therefore, if we want to help both the working class today and the poor tomorrow, we should add an aggressive metropolitan-wide land development program to our job access and public transportation program.

What I foresee as desperately needed is a public purchase of land all over the metropolitan region for public site development—site planning, utilities, schools, recreation, clinics and medical centers, shopping—all should be provided with the grading of the land. Then the land could be sold at reduced costs, as in urban renewal, to private home builders who could be charged with constructing houses in working-class price ranges with or without further subsidies. Under such a program the private builders could once more enter the massive working-class housing market in a permanently useful way, the poor would benefit immediately from the increase in the supply of housing, and there would be something good to trickle down in later years.

ECONOMICS IN THE INDUSTRIAL STATE: SCIENCE AND SEDATIVE ECONOMICS AS A SYSTEM OF BELIEF

By JOHN KENNETH GALBRAITH
Harvard University

I

A recurring and not unsubstantiated charge against economics over the last century has been its employment, not as a science, but as supporting faith. In this latter role it is held to serve not the understanding of economic phenomena but the exclusion of lines of thought that are hostile or unsettling to the discipline or, a related matter, to an influential economic or political community. "Economists" Marx described as "the scientific representatives of the bourgeois class,"¹ and he held that after the bourgeoisie conquered power in England "it was no longer a question" for political economy "whether this theorem or that was true, but whether it was useful to capital or harmful, expedient or inexpedient, politically dangerous or not. In place of disinterested enquirers there were hired prize-fighters; in place of genuine scientific research, the bad conscience and the evil intent of the apologetic."² Veblen, after saying that the competitive model of classical economists "affords the test of absolute economic truth," went on to assert that "the standpoint so gained selectively guides the attention of the classical writers in their observation and apprehension of facts" ³ Tawney observed that during most of the last century the conflict between "individual rights and social functions was marked by the doctrine of the *inevitable harmony* [my italics] between private interests and public good."⁴

This view of economics is not confined to the great dissenters. There would now be considerable agreement that economic theory before Sraffa, Chamberlin, and Robinson excluded from consideration market structures which could not readily

be reconciled with the competitive model or the limiting case of single-firm monopoly. This affirmed a view of economic society in which firms (by implication small) were numerous in the market and without market power and in which the tendency was to an equilibrium of normal profits and optimal resource allocation. This in the United States was over a period—say from 1880 to 1930—when industrial firms were becoming very large and, by all outward sign, wielding great market and political power. In denying scientific recognition or legitimacy to this trend economic theory was not being politically and socially neutral. It was persuading its communicants to avert their eyes from reality. Except where monopoly or intent to monopolize could be shown, the theory denied the need for any social response to economic power. It was playing an active—an actively conservative—role in the political process.

The social and political role of economic belief was at least equally great in the case of Say's law of markets. We now marvel at the hold exerted by this proposition on economic thought before Keynes. And the practical and political consequences (again conservative) were equally profound. If there could be no deficiency or excess in aggregate demand (if any other solution meant that a man was unlearned in the fundamentals of economics)⁵ there could be no case for increasing or decreasing public outlays or revenues to affect the level of output or employment. The alternative possibilities allowed only for a self-correcting theory of the business cycle or one that permitted (or encouraged) the adjustment, i.e., reduction, of wage levels or the correction of other special equilibrium error. On avowedly scientific grounds the discipline thus helped to exclude from consideration what are now commonplace measures of fiscal policy and, *pari passu*, to defend a minimal role for the state. This was accomplished by a

¹ Karl Marx, *The Poverty of Philosophy*, Chap. II (1847).

² Karl Marx, *Capital* (author's preface to the second edition, 1873).

³ "The Place of Science in Modern Civilization: The Preconceptions of Economic Science," in *What Veblen Taught* (Viking Press, 1947), p. 111.*

⁴ R. H. Tawney, *The Acquisitive Society* (Harcourt, Brace and Co., 1920), p. 27. He added that it was the further achievement of economics that "competition was an effective substitute for honesty."

⁵ To be consigned, as Keynes suggested, to "live furtively, below the surface, in the underworlds of Karl Marx, Silvio Gesell or Major Douglas." John Maynard Keynes, *The General Theory of Employment, Interest and Money* (Harcourt, Brace and Co., 1936), p. 32.

proposition which, in the context of the modern industrial economy, virtually all economic scholarship holds to be wrong and even derisory.

One further aspect of this history is important. Popular perception of the shortcoming ran well ahead of the theoretical economic accommodation. While economic theory had no appreciable reaction to the rise of the great industrial firm prior to the 1930's, the case of single-firm monopoly apart, the ubiquity and omnipotence of "big business" had been a source of popular discussion and concern for forty years. It was the basic fare of the muckrakers and the political base of the populists. Journalists and politicians and the public at large had sensed what the theory denied or ignored; namely, that where the participants in an industry were large and few they wielded great power not explained by the occasional case of single-firm monopoly. Similarly, long before Keynes made it reputable for economists, the lesser breeds without the discipline—politicians, journalists, liberal businessmen as well as Gesell, Major Douglas, Foster and Catchings and the other members of the pre-Keynesian underground—had argued that in depression affirmative action should be taken by the state to increase aggregate demand.⁶ A not wholly irrelevant consequence of the rigid and enduring commitment to Say's law was that the economics profession, through the early years of the Great Depression—indeed until rescued by Keynes—had a reputation for doctrinaire negativism. And those who continued to find truth only in the established belief were doomed to live out their lives in a state of obsolescence that was all too cruelly manifest and which, one trusts, will be a sobering lesson for the future.

In yet other instances economics had excluded socially inconvenient analyses, at least until some combination of pressure—the need for practical action, the social intuition of the nonprofessional, competent heresy within the profession—has upset the accepted view.⁷ But I am not con-

cerned with making a catalogue. I wish to argue that present professional belief—the neoclassical model of economic process—as profoundly accepted as was once the competitive model or Say's law, is now similarly excluding urgent as well as politically disturbing questions from professional economic vision. It is important that all be reminded that there is nothing novel about this. On the contrary, it is quite normal—a commonplace aspect of the sociology of the discipline. So, also, is vehement insistence that economics is wholly scientific and neutral when it is being politically quite purposeful. Say's law was most indignantly asserted as a test of professional respectability in the years just before it demise. It was then that it most needed energetic defense. But let me summarize. The accepted economic models, in the past, have not necessarily been the ones that illuminated reality. They have frequently served to divert attention from questions of great social urgency which, in the established view, had alarming implications for political action. In doing this, economics has served a political function. It has been not a science but a conservatively useful system of belief defending that belief as a science. And knowing, and indeed agreeing, that this has occurred before, our minds must be open (or less incautiously closed) to the possibility that it may happen again.

II

The assumption that economics must now abandon, subject to some later definition, is that of consumer sovereignty—and, in light of the role of the modern state in the economy, what might also be called "citizen sovereignty." If this is not done, the discipline will serve, indeed is now serving, not as an elucidation of social phenomena but as a design for suppressing inconvenient social conclusions and action. And given the pressure of present circumstance, that of popular intuition and (one trusts) the growth of intradisciplinary dissent, it will not so serve for very long. My in-

⁶ It was, one senses, the desperation bred of the Great Depression and the willingness so induced to look anew at old truths, as much as the cogency of Keynes's argument, which led to the rejection of Say's law. Before *The General Theory*, liberal reformers such as Paul Douglas in the United States and William H. Beveridge in Britain prescribed for the depression within the framework of Say's law. Not budget deficits to expand demand but wage reductions to increase employability were urged. In the United States both Presidents Hoover and Roosevelt (and also the Hearst press) had embraced a policy of either tax reduction or public spending to raise the level of aggregate demand before Keynes made the idea generally acceptable to economists.

⁷ Including the commitment of the theory of the

firm to the entrepreneur who combines ownership with direction of the business enterprise. With this goes a greatly strengthened commitment to profit maximization as a goal, a determinant market response in pursuit of that profit and an effective exclusion from consideration of other social and political constraints by the corporation on its participants or the public. On this see R. A. Gordon, *Business Leadership in the Large Corporation* (Brookings Institution, 1945), p. 11 *et seq.* Also Robin Marris, *The Economic Theory of 'Managerial' Capitalism* (Free Press of Glencoe, 1964), p. 62 *et seq.* I discuss this at some length in *The New Industrial State* (Houghton Mifflin Co., 1967), Chap. X.

tention in this paper is to put the case for, and consequences of, the changed assumption in the shortest form consistent with necessary qualification and technical precision of argument.⁸

There are three plausible views of the individual in economic society of which two are broadly consistent with the neoclassical model. In the first the individual is regarded neutrally as a participant in a process for transmitting change. The change may be autonomous with the individual—a change in taste reflecting some change in his life design—and its effects are then transmitted through the market to the producer. Or change may originate with the producer, e.g., some change in the production function arising from spontaneous technical innovation, and it is transmitted through the market to the individual. In each case there may be secondary or tertiary reverberations. In each concern is with the process;

⁸ The surrender of the sovereignty of the individual to the producer or producing organization is the theme, explicit or implicit, of two books, *The Affluent Society* (2nd ed. rev., Houghton Mifflin Co., 1969) and *The New Industrial State* (Houghton Mifflin Co., 1967). In both of these books I faced the problem of discarding ideas, much beloved, that had long been part of my habit of thought and also the terrible tendency to recoil when one's analysis suggests or seems to imply practical action well outside the accepted modalities. I was also, as I have said before, faced with the peculiar problem of persuasion that is here involved. A scientific proposition is refuted by proof to the contrary. Belief, especially if it is playing a functionally protective role in the society, is by no means so vulnerable. The strategy of persuasion thus required, as I have also elsewhere made clear, repays some thought. All social disciplines, and perhaps especially economics, are naturally jealous of the larger framework of assumptions in which they operate. For if assumptions become obsolete, so does the knowledge subtended thereon. This vested interest is further reinforced by the functional role of the ideas in excluding inimical lines of thought and action. It follows that to attack such a framework of assumption from within the discipline is a perilous matter. The jury, or most of it, is a party at interest. The fate of all who attacked Say before Keynes is a warning. The alternative is to engage a larger public and thus, as it were, force the issue on the discipline. For, if the assumptions being attacked are vulnerable—if they are incongruent with reality—the public intuition will be responsive. So will be that of the social radical. And if enough such support can be enlisted, the old framework can be broken. The use of this technique naturally incurs a certain measure of professional discomfort. It bypasses the system by which ideas and innovations are submitted for professional scrutiny and winnowing before being passed along to students and the lay public. And it similarly renders nugatory the process by which the intellectual vested interest is protected. To the legitimate rebuke for the first is added the more personal discontent inspired by the second.

no special assumption is made as to the source of the change or the purpose of the process. It should be noted that all changes are transmitted through the market; there is no significant extramarket process by which the producer is brought to accept changes sought by the consumer or by which the consumer is conditioned to accept changes sought by the producer. Most modern mathematical models of microeconomic relationships are, broadly speaking, of this kind.⁹ Public goods are not very satisfactorily embraced by this model.

The second possibility involves a substantial measure of implicit theorizing. The view is still of a process. The process is still a neutral transmitter of change including that originating with the producer. But the ultimate guidance is seen very clearly as coming from the individual; it is to him that the ultimate accommodation is made. The accommodation to changes in the producer's cost function is neutral and technical; the accommodation to changes in the consumer's demand function is functional and moral. It embodies the purpose of the system. Borrowing from political theory a similar though less precise accommodation is made to the changing preferences of the individual citizen-voter for public goods.

None of this need be absolute. The consumer is admitted to be subject to influences that are external to the market. Some of these originate with the producer or the process by which he is supplied. These include specific persuasion by the producer, the more general effect of the cultural emphasis on goods and the competitive and emulative influences which bear on consumption and which, as Professor Duesenberry pointed out many years ago, associate consumption with success in life and thus make it an end in itself.¹⁰ And for private—and, even more especially, public—goods information is transmitted imperfectly by the market. In consequence, welfare economics concerns itself with how the process can be corrected and the consumer equilibrium be made to serve more precisely the individual's preference for kinds and quantities of goods. However, both the extramarket effects and the shortcomings are peripheral; one concedes them in order to protect the larger fact. That larger fact is the ultimate accommodation of the economic system to an individual choice that is original and innate. That accommodation is inhibited and diverted and

⁹ I am grateful for suggestions here from my colleague, Leonid Hurwicz.

¹⁰ James S. Duesenberry, *Income, Saving and the Theory of Consumer Behavior* (Harvard Univ. Press, 1949), p. 28.

modified but only as the brush along the banks and the rocks along the bottom inhibit and divert and modify the flow of a stream.

This accommodation, it should specifically be noted, is broadly consistent with the accepted theory of monopoly or oligopoly. The demand function of the individual is given, which is to say that it originates independently of the producer. The producer seeks to maximize revenues—a vital point. Changes in individual demand when aggregated lead accordingly to responses that are no less reliable than those in the competitive market. The resulting distribution of resources and income is different and so is the resulting consumption. By welfare standards it falls short of an ideal. But it is not different in being less responsive to the ultimate authority of the consumer.

The third possible view sees the process as one in which the ultimate accommodation in a significant part of the economy is to the producer. The individual's wants, though superficially they may seem to originate with him, are ultimately at the behest of the mechanism that supplies them. In the most specific manifestation, the producing firm controls its own prices in the market and goes beyond to persuade the consumer to the appropriate responding behavior. But it also selects and designs products with a view to what can be so priced and made subject to such persuasion. And it does this in a society in which the strongly iterated and reiterated praise of goods makes them seem important for happiness and thus makes the individual attentive to claims in this regard. And the persuasion proceeds in the context of a generally affluent supply of goods, which means that their contribution is to psychic rather than to physical need. The further consequence is that the individual is open to persuasion—to appeals to his psyche—as he would not be were physical effects alone involved.¹¹ On occasion the

¹¹ Some will be aware of the energy with which I have pressed this distinction. Cf. *The Affluent Society* (2nd ed., *op. cit.*, p. 134 *et seq.*). It is one of those naively crucial matters (as Keynes earlier held) on which much turns. Economics generally denies the distinction between physical need and psychic satisfactions—taking advantage in part of the undeniable fact that the line between the two lends itself to no precise conceptual demarcation. Thus, it excludes from consideration the notion of a class of wants which, originating in the psyche, are subject to management by psychological means as wants originating in physical need are not. This greatly defends the values of a society which measures achievement by output. There being no valid difference in the wants being served there is no lessening of the urgency of output. The notion of production for frivolous purposes is almost completely elided. Thus, the

state will supply ancillary services that are needed to obtain the sought-after behavior of the individual—the provision of highways as an aspect of the management of consumer behavior by the automobile industry is an obvious example. By regulating aggregate demand the state also insures that the microeconomic management of demand will not be nullified by macroeconomic movement.¹²

This view of economic process extends with emphasis to public goods. Here for important classes of products and services—weapons systems, space probes and travel, a supersonic transport—decisions are taken by the producers, i.e., the armed services and the supplying firms, in pursuit of their own goals. The Congress and the public are then accommodated or commanded thereto.¹³

The need to manage consumer behavior, as I have argued in detail elsewhere,¹⁴ arises from the circumstances of modern industrial life—sophisticated technology, large commitments of capital, long-time horizons in product development and production and, in consequence, large, inflexible and vulnerable organization. These lead, in turn, to the need to control as many as possible of the parameters (costs, prices, demand, costs and risks

importance of production remains above question. Once again one sees economics overriding a commonsense view to defend what is, unquestionably, a most convenient conclusion.

¹² The one is obviously dependent in a highly practical way on the other and it is a curiosity of economics that the two—the need to insure that people will want G.M. cars and the need to insure that they will be able to buy G.M. cars—has been so little associated.

¹³ The most meaningful distinction between a market and a planned economy, so it seems to me, turns on whether and to what extent accommodation is to producer or consumer choice. The more responsive the producer must be to consumer choice, the more it is a market economy. The greater his power to establish prices and to persuade, command, or otherwise arrange the consumer response at these prices, the more it is a planned economy. Intervention by the state does not alter the fact of planning; it changes only its nature, extent or efficiency. In everyday language, planning means the systematic exercise of foresight. This is a source of ambiguity for even within narrow market parameters there can be exercise of such foresight—specifically to anticipate market behavior or make more efficient the firms' response. James E. Meade, in his review article, "Is *The New Industrial State* Inevitable?" (*Econ. J.*, June, 1968), rightly points out that I do not distinguish adequately between such planning within the market instruction and planning which embraces the decisions of the consumer or citizen.

¹⁴ *The New Industrial State*, *op. cit.*, pp. 1-97.

of technological innovation) within which the firm operates. This development is greatly different in different parts of the economy—the range is from the producer of modern weapons systems or automobiles at one extreme to the vegetable farmer or small shopkeeper at the other. The extent of the accommodation of the individual to producer need varies accordingly. This difference is not something to be minimized; on the contrary, it is itself of practical consequence for economic behavior, as I will argue in a moment. The efficacy of the management of the consumer or the public in any industry will also vary over time and, on occasion, will partly fail or be frustrated. This management is exercised at increasing cost which varies as between products and market structures.¹⁵

Maximization in this model is not of profits alone but of the panoply of organization interests—security and autonomy of the organization, growth (and consequent increase in pay and opportunity), technical achievement, public prestige, as well as profits. The priority accorded the several goals will plausibly differ somewhat for different organizations.

Finally, it remains possible, at least in the private sector of the economy, for the individual to contract out or partially out of the management to which he is subject. (This, more than incidentally, may allow him to deny the existence of such management and to point to his own immunity as proof. A certain part of the case for unmanaged consumer choice rests subjectively on such grounds.) All of these qualifications are essential for only the inexperienced rejoice those who are resistant to an idea by allowing themselves the catharsis of overstatement.

III

So far as anything in economics is certain, it is that the first two of the foregoing views have a monopoly of established belief. Formal microstatic models emphasize the first view; the less formal, more intuitive and more influential writing and instruction assumes the second. It is not impossible (though not altogether easy) to find work that concedes producer management of consumer taste. Tibor Scitovsky¹⁶ has dealt interestingly with the management of consumer markets on behalf of the majority taste—an argument

¹⁵ Slowly increasing costs of persuasion for (e.g.) automobiles or soap partly distinguish these industries from vertically increasing costs in, say, agriculture.

¹⁶ *Papers on Welfare and Growth* (London: George Allen and Unwin, Ltd., 1964), pp. 241–49.

with more than parenthetical importance for the economics of the arts. Jerome Rothenburg¹⁷ has held of advertising that, although it “is probably not accountable for drastic changes, it is reckless to assume only trivial impact”¹⁸ and noting that there are “endogenous taste changes—changes induced by producer investments designed to effect such changes,” he has concluded that “few would insist that the consumer is sovereign in any useful sense.”¹⁹ A number of other scholars, some in more recent times accepting my arguments, have agreed. But these are exceptions. In the general view economics is a process by which the individual imposes his will on the producer—as put matter of fact by Fisher, Griliches, and Kaysen, “there *is always* an assumption of consumer sovereignty in the market economy.”²⁰ (My italics.) And, although the process is confused, indirect and inefficient, the citizen is equally assumed to impose his ultimate will as to public goods on the state. When one comes to the world of the textbook, an important matter when, as here, one is concerned with economics as it serves functionally through its assumptions to influence belief and thus action, the commitment to consumer (and citizen) sovereignty becomes virtually absolute.²¹

IV

It is not my purpose here to argue that the accepted views are incognate with reality, that the

¹⁷ “Consumer’s Sovereignty Revisited and the Hospitality of Freedom of Choice,” *A.E.R.*, May, 1962.

¹⁸ *Ibid.*, p. 280.

¹⁹ *Ibid.*, p. 279.

²⁰ Franklin M. Fisher, Zvi Griliches, and Carl Kaysen, “The Costs of Automobile Model Changes Since 1949,” *J.P.E.*, Oct., 1962, p. 434.

²¹ “What things will be produced is determined by the votes of the consumers—not every two years at the polls but every day in their decision to purchase this item and not that.” Paul Samuelson, *Economics* (7th ed., McGraw-Hill Book Co.), p. 42. However, in this edition, Professor Samuelson subsequently softens this proposition and I sense, otherwise, that his commitment to consumer sovereignty is far from rigid. Others are more categorical. “. . . only [the consumer] can make the crucial decision on what goods he most prefers; thus, in the final analysis, consumers collectively decide what industry is to produce. The choices of consumers provide the basis on which business makes its decisions.” C. E. Ferguson and Juanita M. Kreps, *Principles of Economics* (2nd ed., Holt, Rinehart and Winston, 1965), p. 80. “As buyers, individually but totaling millions react to prices, they also change prices. Consumers vote with their dollars. The buyer, himself guided by relative prices in making his choices, is directing the allocation of productive resources.” C. Lowell Harris, *The American Economy* (4th ed., Richard D. Irwin, Inc., 1962), p. 380.

third view is right. This I have done at length elsewhere.²² It is hard to believe that the uncommitted reader will find the case for producer sovereignty in the form in which I have just outlined it wholly implausible. The case is perhaps strongest for public goods; there can be few men of available mind who have recently looked at the process by which the national defense is provided without wondering if the conventional view of ultimate citizen sovereignty is acceptable. This is not a detail; it is half the federal budget. And many must have wondered if the conventional view, indistinct as it is, might not be serving to divert attention from the disenchanting reality—if it did not accord the public the mythology of power while giving the military bureaucracy and associated industries the reality of power. But in the large-scale consumers goods industries the case is not greatly less convincing. There is the massive outlay on persuading the consumer.²³ There is also the general increase in the persuasive effort paralleling the development of increasingly complex technology and organization. It is a tenet of the more developed consumers goods industries that products must be selected, designed, and produced with a view to what lends itself to persuasion. Accordingly, it involves an exercise of imagi-

nation to suppose that the taste so expressed originates with the consumer. What the consumer deems to be a desirably shaped and chromatically compelling automobile is substantially different this year from what it was five years ago. But few would wish to argue that this represented a change in the consumer's intrinsic and improving vision of a vehicle—that, indeed, it was other than something accomplished with no slight skill, art and expense by the automobile producers. It is not necessary to argue that the management of the consumer by the producer is complete, only that it makes consumer behavior conform in broad contours to producer need and intent. This is plausibly in accord with everyday observation of marketing practice and the commonplace claims of its practitioners.²⁴ Nor will many resist the idea that these industries can bring the state to the support of their efforts in creating and managing consumer wants—that the automobile companies can get the highways that are essential for a consumer preference for automobile transportation; that the airline and aircraft manufacturing companies can win public financing for the development of new types of aircraft, in the past under military guise but now quite overtly in the case of the SST; and that the tobacco companies can obtain extensive governmental immunity from the scientific evidence on the causes of cancer.

Finally, few will doubt the enormous stress which the process of persuasion places upon the importance of goods and the belief so created of the nexus between goods (including those that are technically innovative or can be so represented) and happiness. This, most will suppose, increases the susceptibility of consumers to persuasion. If goods are firmly established as the cause of happiness, the public will be attentive and responsive to claims to reward on their behalf, and certainly

²² *The Affluent Society*, *op. cit.*, especially pp. 134-67 and *The New Industrial State*, *op. cit.*, especially pp. 159-218.

²³ There is a marked tendency, especially among the unconsciously tendentious defenders of the market and thus of consumer and citizen sovereignty, to denigrate and even dismiss the role of advertising. One recent critic disposes of my interest in it by saying that it is concerned with "the most hackneyed theme in modern social literature—the power of advertising." (Scott Gordon, "The Close of the Galbraithian System," *J.P.E.*, July-Aug., 1968, p. 642.) So to minimize the role of so vast, obtrusive, expensive, and integral an aspect of the modern market must surely provoke question. One notes also that advertising has continued to be a somewhat indigestible lump in conventional microeconomic theory. To see it, as does the most commonly accepted oligopoly theory, essentially as a functionless but safe alternative to price competition which ultimately cancels itself out, is not altogether satisfying and leads inevitably to the question, ill-received by advertisers and media when not tactfully elided by economists, as to why such a portentous waste is not prohibited or mightily taxed. But there is also the fact, as Professor Rothenburg points out, that advertising is the most direct and visible attack on the concept of consumer sovereignty. This, one at least suspects, may be a reason for wanting to ignore it or, failing that, to follow Professor Gordon in suggesting that concern with it is unfashionable or otherwise intellectually unworthy. I count it an important part of the case for producer sovereignty that its exercise gives to so important an activity as advertising a wholly functional role in economic life.

²⁴ Consumer management is a more complex business where, as in the characteristic oligopoly case, a few large firms produce a closely substitutable product. Here predictability of consumer behavior is enhanced by the management of consumer taste and, of course, reduced by the fact that others are seeking to do the same thing with greater or less effect. However, as I have elsewhere argued (*The New Industrial State*, *op. cit.*, p. 206 *et seq.*), from the aggregate of this effect—the general attraction to the common products of the industry and the success of one firm, the inevitability that it ride with success, the stimulated response of others—comes an equilibrium more predictable for any fully participant firm than would result from unmanaged demand. And, as I note above, there are further important effects from this process and the effort expended upon it on the general social attitudes toward goods and their producers.

the relentless propaganda on behalf of goods must greatly increase the importance attached to production. This, in turn, strengthens the position of producers in the exercise of their sovereignty especially as regards the community and the state. What can be so important as what they do? Economics again assists by making the level of output the formal, measurable accomplishment of the society. But my purpose is not to argue the case for producer sovereignty but to assume it (though less comprehensively) as consumer sovereignty is now assumed. And assuming producer sovereignty I want to look at the features of the society which, excluded from view by the assumption of consumer sovereignty, then swim almost majestically into view. What is so solved makes my case.

V

The first and by far the most important matter that thus becomes clear concerns the relation of the individual to industrial society in the largest sense. In the accepted economics, no general conflict can arise here. The individual or citizen is sovereign. There may be differences between different individuals as to whose commands are heeded. By ancient classical assent, the rich speak more authoritatively in markets than the poor. And there may be friction or aberrations in the response of institutions to the ultimate authority of the individual. But none of this is systemic. The individual is ultimately and fundamentally in command; he cannot be at war with himself.

When producer sovereignty is assumed, the result is very different. This sovereignty is exercised, we have seen, by large and complex organizations. This exercise of power is to serve their own goals—goals that include the security of the organization and its growth, convenience, prestige, commitment to technological virtuosity as well as its profits. There is every probability that these goals will differ from the aggregate expression of individual goals. Individuals are then accommodated to these goals, not the reverse. This normally will involve persuasion. But it may involve resort to the state or, in the manner of a utility marching its lines across the countryside, to power that is inherent in institutional position.

The consequence of economic development, so viewed, is not of harmony between the individual and economic institutions but of conflict. The conflict is modified by the persuasion—but not for the unpersuaded or those who sense what is happening. This conflict is sharply at odds with accepted economic (and political) interpretations of the reality. But it is not at odds with the reality. If there is an agreed diagnosis of contempo-

rary discontent both in the United States and the other industrial societies, it is that the individual feels himself in the grip of large, impersonal forces whose purposes he senses to be hostile and in relation to which he feels helpless. The Pentagon pursues wars and builds weapons systems in accordance with an inner dynamic. Similarly NASA. So the Department of Transportation in relation to the SST. So General Motors as a producer of automobiles that threaten to smother cities and as a sponsor of highways that have already gone far to devour them. So industry generally as it subsumes countryside, water and air.

This conflict comes to a peculiarly sharp focus in the universities. This also is what the model would lead one to expect. In the universities large numbers of students are brought together by the unprecedented demands of the industrial system for qualified manpower. They are given a sense of personality as the older industrial proletariat was not; the older proletariat, indeed, was taught by the unions to submerge personality into a sense of class. And students are also exposed with some sense of righteousness to social doctrine—economic and political theory—which holds that the individual is possessed of ultimate power. And, in contrast, they see a world in which organization exercises large, even seemingly plenary power and to which they, as citizens, soldiers, consumers or organization men are expected to be subordinate. None of the resulting discontent could occur in a society in which the consumer or the citizen is sovereign. It is surely probable, even predictable in a society in which producing organizations are sovereign—in which they have power to pursue purposes of their own that are different from those of the consumer or citizen.

The notion of producer sovereignty, then, is not only empirically plausible—a seemingly logical response to the needs of the modern, highly technical, highly capitalized, very complex industrial organization—but it also sharply illuminates our major present concern. This is a good thing for any social theory to do. But economic and associated political theory in remaining with the notion of consumer and citizen sovereignty are not merely failing to interpret reality. By contributing to a contrast between what is taught and what exists they are weakening confidence in the objectivity of social science—and perhaps even in education itself. They are making these the servant not of an understanding of reality but of a conservatively useful myth that conceals the reality. But since, in fact, this cannot be concealed they are adding to frustration and conflict.

But this is not all. In other respects the notion

of consumer and citizen sovereignty is diverting attention from fundamental problems of the economic and political system in a fashion that serves to strengthen the very producer sovereignty that the discipline denies. Let me cite eight specific examples, each of them of no slight contemporary concern.

If the mix of goods being produced at any given time seems unsatisfactory—if there are too many automobiles, too little mass transport—consumer sovereignty holds that this reflects the dominant consumer will. Similarly, if housing is scarce and poor, housing appliances abundant and efficient. The person who expresses doubt is seeking, in undemocratic elitist fashion, to substitute his taste for that of a majority. But if producer sovereignty is assumed the product mix will be the expression of its comparative power. If there appear to be too many automobiles, insufficient intercity or commuter rail service, or urban rapid transit, this will plausibly be because the automobile industry exercises its sovereignty (including its power to persuade people that they want automobiles more) more effectively than do the producers of alternative transport. We have more appliances than houses because General Electric is more powerfully sovereign than the house builders. Except to the exceptionally devout, none of this, I venture to think, will seem unreasonable. But economics as it is taught, by emphasizing consumer sovereignty, makes itself a shield for the exercise of producer sovereignty by the automobile industry. For by making questions about too many automobiles an elitist and undemocratic interference with consumer choice, it effectively excludes questions about the power of the automobile industry to impose its preference. In effect, it gives high moral sanction to social indifference.

The concept of consumer sovereignty acts with marked force to inhibit questions concerning the cultural achievements of the system. It will surely be agreed that whatever the effects of advertising its ultimate effect is an extremely powerful and sustained propaganda on the importance of goods. No similar case is made on behalf of artistic, educational, or other humane achievement. The notion of consumer sovereignty suppresses the response.²⁵ While it may be conceded that the pop-

ular taste is biased toward goods, it insists that the popular taste be respected. The notion of producer sovereignty, by contrast, forces recognition of the inconvenient certainty that the source of much of the taste is in the producing organizations that promulgate it for the community. Economics renders a further conservative service. To the microeconomic doctrine of consumer sovereignty it adds the macroeconomic test of output not art as the measure of social achievement.

The concept of consumer and citizen sovereignty allows of no organic likelihood of a bias in the economy for private as opposed to those public goods that do not serve producer sovereignty. At most, there will be blockages and error in the allocation of resources to the public sector. Producer sovereignty, coupled with the fact that the instruments of its exercise, advertising for example, are elaborately and expensively available to the private economy and not available in any similar fashion to the public sector, makes this bias systemic. In these past weeks the United States Senate has been hastening to reduce taxes in face of the seemingly unprecedented need of the civilian public services. And in the background has been the doctrine that, unprecedented private consumption notwithstanding, taxes now bear on people with unprecedented weight. Something must surely be attributed to the superior ability of producer sovereignty to persuade as to the urgency of private goods. If this be agreed, then again it cannot be entirely bad to have a theory that explains the contemporary reality.

Consumer and citizen sovereignty sanctions the current claims on resources of the military industrial power—it is in response to the perceived need and expressed demand of the public. Or, alternatively, it is a *sui generis* error—a fault in an otherwise workable system. The first view will tax belief of even the most committed supporters of the received model; the second, as an explanation for any part of the economy that is so large in both claim on resources and social portent, must lack something in scientific appeal. The notion of producer sovereignty increasing in effect with increasingly complex organization and technology brings the power of the producers of military goods and services wholly into focus.

Consumer sovereignty makes pollution and other environmental disharmony an external diseconomy. The cost of damage to air, water, and

²⁵ With a peculiarly righteous indignation, in fact. I made this case in less sharp form and with much stronger emphasis on public goods in *The Affluent Society*. The rebuke differed only in emphasis. A few held that I was presuming to set an admittedly attractive judgment against the democratic manifestation of the market. The rest held that I was presum-

ing to interpose a precious, narcissistic, arrogant or otherwise grossly pretentious judgment for that of the market.

surroundings is borne by the community, not by the producer. Since the market is assumed to be an efficient expression of public taste and need, external diseconomies have long been viewed as of peripheral significance to be corrected by essentially cosmetic public action. With producer sovereignty environmental damage becomes a normal consequence of the conflict between the goals of the producing firm and those of the public. Its particular focus is the emphasis which the firm places on growth and freedom for its organization for autonomous decision unhampered by community or public constraints. Here again economic theory in its macroeconomic norm strongly supports producer sovereignty. It powerfully supports the argument, commonplace in these matters, that nothing and certainly not the minor ecological preferences of the community, should interfere with the stern needs of production and productive efficiency. These give the power line or the industrial effluent a natural priority.

Consumer sovereignty allows of no question as to a socially desirable upper limit to the consumption either in general or in particular products. The consumer wants more; theirs not to reason why, theirs but to satisfy that want. With producer sovereignty the level of consumption is seen to be a derivative of producer goals including the producer commitment to growth. Consumer attitudes are seen to be substantially formed by producer persuasion on behalf of goods. The question must then arise as to whether General Motors is the proper agency to decide the proper level of consumption for its products. And since the matter is not decided by the collective inner will of the public the question also arises as to the optimal upper level of production and consumption in general. This question should, perhaps, have been faced before now. For a host of reasons, including the effect on environment, it is unlikely that we can continue to increase physical output at recent past rates for the next (say) twenty years. It is easy to see how, once again, economics has rendered conservative service. By holding this matter to be resolved by the inner collective will of the public, it has effectively banned from public discussion all question as to how much a community should produce or consume.

In the conventional model differences in income for personal services reflect ultimately the willingness of the community to pay for such services as derived from market desires and preferences. Inequality in nonproperty incomes thus derives a substantial measure of functional sanction—a not unimportant matter at a time when the inequality of income distribution is increas-

ing.²⁶ Producer sovereignty makes this income inequality, at least in part, the product of bureaucratic design, tradition and self-arrangement. Such a cause of inequality enjoys no similarly high sanction. It does correspond, however, with the everyday appreciation of the matter by the participants.

Finally, a more immediate point. If consumer sovereignty is assumed, there will be a strong presumption that actions directly or indirectly affecting the consumer's market behavior will have a strong and reliable market response. It is to the consumer that the market responds. If by either fiscal or monetary policy his outlays are directly or ultimately curtailed, there will be confidence in the ensuing effect on prices and production. With producer sovereignty there will be no similar confidence. The producing firm is pursuing its preferred goals which is to say that it is maximizing not necessarily its profits but its organizational interests. If this has caused it to subordinate profit maximization to growth, it can, if it must, increase revenues by increasing prices. And its organizational interests will include the security of the organization as opposed to the dangers inherent in labor conflict or interrupted production. So, given producer sovereignty, it is quite predictable that efforts to limit consumer expenditure in an inflationary context, even if successful, will be accompanied by continuing price and wage increases in the highly industrialized, highly organized sector of the economy. The fact that this sector is not coordinate with the whole economy is of especial importance here. It means that the part of the economy characterized by producer sovereignty in effect exports its tensions to the more vulnerable sector where consumer sovereignty is still relevant.²⁷ A measure of index stabilization may even be accomplished at the expense of the latter.

I say that this is a more immediate point. It is an unduly succinct but wholly accurate descrip-

* Joseph A. Pechman, "The Rich, The Poor and The Taxes They Pay," *The Public Interest*, Fall, 1968, pp. 21-43.

²⁷ In *The New Industrial State*, influenced by the comparative success in the first half of the 1960's, in stabilizing prices in the organized sector of the economy through the guideposts and by the parallel resort of numerous other industrial countries to some form of wage-price restraint, I concluded that this was one of the parameters (like minimum prices or stable aggregate demand) where large organization would accept and even seek public stabilizing action. I still think public opposition to inflation as well as balance of payments and other reasons will eventually force such action. I am no longer so certain that it is one of the things that large organization needs.

tion of what is now happening. And here accepted economic theory serves not only to divert public attention from requisite action—the replacement of the sovereign producer with the sovereign state in the process of price determination—but it clearly blinds the eyes of the economists who are responsible for policy. In consequence, month after month, they continue optimistically to avow their hope and intention of ending inflation by measures appropriate to consumer sovereignty. And not surprisingly, month after month, they are roundly defeated by a reality reflecting producer sovereignty.²⁸ In an age when public officials are often thought averse to personal sacrifice or immolation in pursuit of principle, it is gratifying in a way to find that economists are still willing to surrender their professional reputations on the altar of established doctrine. One regrets only that it is not in a more useful cause.

VI

None will doubt that this paper leaves many important questions unanswered. There is, notably, the question as to the theory of the state that is here implied. The state as here envisaged comes close to being the executive committee of the large producing organization—of the technostucture. It stabilizes aggregate demand, underwrites or socializes expensive or risky technology, reflects the will of large organization in the mix of military and nonmilitary public goods, provides such needed public artifacts as highways for the management of specific consumer demand, supplies qualified manpower, otherwise stabilizes those parameters or does that planning which the large producing organization cannot do for itself. This being so, one must ask if the industrial state can separate itself from organization—if it can be the instrument of individual will. Let us not imagine that it will be easy.

One must ask also if there is a choice or a trade-off between increased technology, increased complexity of organization, and increased production on the one hand and increased power of individual expression on the other. If so, is there a substantial measure of social perception in the behavior of the young who (at least while young) see in the rejection of physical artifacts an avenue to greater self-expression?

One must ask further if there is a possibility of meeting the power of organization with the power of anti-organization. If the automobile industry is sovereign in the market and thus in its decisions

on automobile population and their effect on environment, can it be made less sovereign by countering organization—by organization to exclude the internal combustion engine from urban areas? If the weapons industry is sovereign in the Congress, can it be made less sovereign by countering organization which removes its servants from Congress?

Finally, and of high interest for this paper, what is the effect on economics as a discipline, after years of comfortable coexistence with industrial and associated public bureaucracy, if it makes exercise of power by such organization in its own behalf an accepted and central preoccupation? What happens when it views the mix of products, the level of production, the autonomous exercise of power by the weapons industries, the effect on the environment not to mention the resolution of the wage-price bargain as an exercise of bureaucratic power in the interest of bureaucratic goals and not as a reliable if sometimes obstructed response to the ultimate consumer will? Can economic theory embrace such issues? Can it stand up to the resulting contention? Clearly these matters have consequences for economics, as for the society at large, that are not slight. They present an interesting choice for our discipline. Economics can remain with consumer sovereignty and be comfortable, noncontroversial, increasingly sophisticated in its models and increasingly, and perhaps even dramatically, unrelated to life. Or it can accept the implications of producer power—of the sovereignty of the great organizations. Then it will be contentious, politically perilous and for a long while, perhaps, intellectually inelegant in its models. But it will in compensation be relevant to the most immediate and formidable concerns of the industrial society.

I have little doubt as to the choice. Among my generation it will be, in principle, for comfort and its associated refinements. We have had one revolution; Keynes was enough. There are elements of truth in this model, it will be said, but nothing that should require one to change his mind or his pedagogy. I say this will be the choice in principle for it will not be so in fact. Mention of Keynes reminds us that he stressed the ultimate power of ideas. In degree, he was right. But he could wisely have stressed the far greater authority of circumstance. Circumstance has given us the great private and associated public organizations. They have great and evident power. Divorced of this circumstance—as an abstract model interesting for itself—the ideas I am urging here would be nothing. Reinforced by such circumstance they are ineluctable.

²⁸ Until, quite possibly, they achieve stabilization, as previously noted, at the cost of the nonindustrial and vulnerable sector of the economy.

DISCUSSION

ELI GOLDSTON: Some weeks ago, my neighbor and your President, Professor Leontief, asked me if I would come down to New York today to your convention to disagree with Professor Galbraith. I asked Professor Leontief, "What's Galbraith going to say?" and he just sort of shrugged. So I guess that my function is to select from the rich table before us the particular items that don't suit my taste. My qualifications for this are, in a sense, unique. As Professor Galbraith once said in introducing me to some people: "Eli Goldston is best known as having been my tenant while I was Ambassador to India."

Now the reason for my renting the Galbraith home was that I understood he had an extensive library. There I found all of the early Galbraithian books and while he was, according to our Chairlady, studying fertilizer in India, I was studying vintage Galbraith in Cambridge, Massachusetts. I am not suggesting that the concurrence in time of these events has any particular significance, but I do find myself today, like Ken's OPA critics to whom he has unkindly referred, primarily armed with his discarded weapons—the vintage Galbraith. I find that his abandoned doctrines are the best answer to his newest theories.

But being a businessman myself, I guess that I am here cast in the role of an articulate fish at a convention for ichthyologists and that I should approach this role with some humility. However, as Ken has so frequently said: "Humility is an excessively overpraised virtue." So I shall proceed directly with my assignment as given by your President, Professor Leontief, and point out where, I think, not only Ken but those who disagree with him as well are somewhat wrong.

The points that I will try to make are as follows: First of all, producer sovereignty is vastly overstated. While you ichthyologists are looking down from the surface of the water at that coral reef, there are an awful lot of fellows like myself swimming around amongst the other fish, including some of the sharks, managing to survive and seeing quite a different picture than you see from up there on the surface. Second, I would suggest that the materialism and the bad priorities that affect our society today have wider and perhaps deeper causes than some group of industrial demons who somehow have acquired sovereignty over our society. This leads me then to a somewhat different conclusion, which is to recognize that a democracy which is affluent and materialistic has to persuade a majority of voting citizens to give rather than to persuade a majority of voting citizens to take from one minority and give it to another. My analysis, however, brings me to ultimate agreement with Professor Galbraith—a belief that the current obsession of economists with mathematics and the current failure of most economists to link in any fashion the economic interests of today with the vast new learning in the field of social relations has, in fact, made economics into a faith rather than a science. And more than that, economics has been made into a parlor game like

bridge rather than a tool to be used in forging a better society.

The purpose of Professor Galbraith, I believe he has stated, is to affect belief and to effect action. That should be regarded today as the fundamental duty of any professional economist—an interest in affecting a belief in order to effect action—because if we don't take action soon we won't have much of a society left to act upon.

Turn first to the matter of producer sovereignty. Living down there on that coral reef, I simply can't see the world and its workings as a matter of black or white, but rather as a mass of swirling shades of gray changing from time to time, and changing from side to side, changing from day to night. Professor Galbraith treats the lack of success of the Edsel automobile as an exception that proves his rule. He suggests that the change in the appearance of today's automobile—more chromium or larger mufflers or what you will—reflects some foisting by Detroit, on some sort of obedient, compliant fool, of things he never desired. But there are dozens of autos besides the Edsels that have died because the consumer would not comply—witness the DeSoto, or the Bantam, the Corvair or the Valiant or the Falcon. The recent failure of Gablinger Beer—some of you will perhaps recall those insistent commercials for the "low cal" beer—suggests once again that we haven't seen the last Edsel type of new product disaster. Doyle, Dane & Bernbach, the advertising geniuses who supposedly put across the Volkswagen, Polaroid, and Avis—flopped with a product that simply didn't catch the public fancy. Despite the tremendous advertising effort, "low cal" beer turned out to be a hopeless failure. When one reflects on the ratio of exceptions to rules and failures of success in producer sovereignty, the famous line from "Thanatopsis" comes to mind: "All that tread the globe are but a handful to the tribes that slumber in its bosom."

In basing their argument on the appearance of an automobile or the caloric content of beer, the principal difficulty of economists is that they fail to talk to their colleagues at the nearest business school to learn the distinction between marketing and sales. In marketing, you're out there trying to find, as Professor Robin Marris says, how to perceive a latent consumer need and foster it into a felt want. That takes, first of all, a very careful perception of where the customer now is and where he is likely to go and whether or not you can affect his movement by advertising. That decision is made at a point before the capital investments are committed and if you aren't too certain on your marketing you try to make a multipurpose capital investment so that you can tend to change a little bit if perhaps your marketing is wrong. In the marketing part of business efforts you are focusing on nothing but the desire of the consumer, albeit—and this is where some of the confusion comes in—you are considering in a very sophisticated way the degree to

which you can focus that desire on a particular product. The significant thing isn't that people are buying chrome on their cars. The significant thing is that marketing men have learned that consumers desire to buy a car that looks different from last year's car. If you want a good lesson in how the consumer dominates the market, take a look at what happened to the piano industry once the automobile came in. Before automobiles were common, the sign of success was a piano—often not only not played but even without a player in the family. Well, the piano disappeared from the living room as the car went into the garage. The thing that Detroit is responding to is not some forcing down of a new breed of chrome but a compliance with a consumer demand that they be able to display in this materialistic society that they do have this year's car. The amount of irrationality in this marketing study of the consumer can't be underestimated. For example, take the great sales success of Ivory Soap—99.44/100% pure. Pure what? What is the other small percentage? Hydrochloric acid?

When you deal with marketing, you're reaching out into a vast unknown—200 million Americans with irrational and rational, articulated and subliminal desires—but when you come to selling you're in a different kind of business operation. You've already committed the capital; you already have the product on your shelf. That's when you come into the restaurant where there's a big pot of stew in the kitchen and the waitress says to you with a smile: "Try the stew today. It's awfully good." That's selling. Marketing judgment occurred when the cook decided to make that pot of stew because he thought it was going to be very cold and people would come in for hot stew. But when the weather turned hot, he had to tell the waitresses to push the stew. So I would suggest that the extent of producer dominance is confused by a confusion between marketing and selling.

Professor Galbraith talks about pipelines marching across the countryside. Recently I was trying to lay a small underground pipeline through Milton, Massachusetts. We used to do this very easily as Ken suggests, but I think that there is a certain lag between professorial perception of social change and the social change itself. We used to have fellows we paid about \$5,000 a year who were known as "right-of-way" men. They would go out to a housewife and say: "We're going to put a pipeline through your backyard. You sign this and you'll get \$5.00." She'd say: "Aw, gee, I don't want a pipeline in my backyard." He'd say: "What are you going to do, fight City Hall?" Then, five years ago she'd sign. Today she'll say, "You're damned tootin' I'll fight City Hall. I'm gonna go down and take off all my clothes and jump up and down and scream right in the Mayor's office." So nowadays we don't have a \$5,000-a-year fellow roaming around with \$5.00 bills. We have a fellow we have hired away from the State Department and he approaches "the market" with a better understanding of its current interests. The notion of producer dominance, I think, must be vastly modified to recognize what has happened to society just in the last few years as a re-

sult of the examples of the civil rights movement.

Turn then to the question: If producer sovereignty doesn't control, why are we materialistic? Why do we have the wrong priorities? I wouldn't deny that we have today an unfortunate equating of material goods and happiness. But the danger of demonology is that if you kill the wrong demon, or if you focus on one demon when there are many demons, you still won't have solved the problem. I think that in order to understand the attitudes of Americans today you must consider what has happened in all sorts of fields and all sorts of institutions—education, religion, history, psychology, and unions. And although Professor Galbraith says that the consumer cannot be at war with himself, I would say that this is quite possible. Indeed, Professor Kenniston sees such internal conflict as the key to what we oldsters call the problem of today's young people. He maintains that youth's war with himself and with his parents has caused a great deal of the change in youth's attitudes of today. Further, many of our problems today reflect a changed attitude during the childhood years. Anthropologists tell us that drastic changes in the attitudes of society follow relatively dramatic changes in attitudes toward child rearing. If that is so, and unless there is some big industry I don't know about running child rearing, perhaps the problems of our society do go somewhat deeper and somewhat broader than our industrial economy.

This brings us to what I regard as the real problem for society. It is not domination by some materialistic industrial magnates. Our problem is that the affluence of America has changed our income distribution to look more like an egg than a triangle. In the bottom of that egg are the blacks and Puerto Ricans, the Indians and the other minorities and unfortunates of our society. It was very simple when your income distribution was a triangle because the people in the middle could take from the few rich people on top and give to the many poor people on the bottom. But when it is like an egg and you already have a progressive income tax, you must now take from the people in the middle. The rise of a huge middle-income class completely changes the situation. Two-thirds of American families receive incomes between \$5,000 and \$15,000—and almost all pay some federal income tax.

Consider the current tax law to which Professor Galbraith made reference. I think that the disgraceful tax law that was passed reflects very little commercial or industrial pressure. It reflects rather a feeling that the silent majority must be paid off. Let me quote one Senate Democrat: "What we are fighting for is suburbia." Former Budget Director, Charles Schultz, puts it another way: "When the chips are down on tax cuts, those who talked about priorities for pollution control and education and an end to hunger, voted for beer and cosmetics and white wall tires." So I would suggest that what we have to develop is a countervailing political power. We must amass constituencies and remember that they are affluent and literate and taxpayers. We have to do the magnificent job of persuasion that James Webb did with NASA.

It was no accident that its two major installations went to the home state of the then President and the then Vice-President. It was so accident that extensive educational facilities were set up at the prestige universities and as fast as the fellows graduated, funding was provided for other educational institutions carefully geographically scattered so that a vast constituency of political support, of educational support, of popular support was built for the effort to the moon. What has to be done now is for skilled people to amass similar constituencies for social needs. I think George Romney is doing that now in HUD. Secretary Weaver never understood the hard-sell. He had those inspiring program names: Section 235, Section 236, and that most inspiring of all, Section 221 (d) (3). Meanwhile, Sargent Shriver, with a sense of salesmanship, had his Headstart, Upward Bound, Vista. Where does the money go? And the political support? It goes to programs that can attract and continue to build a constituency. Professor Galbraith is right when he says we have more appliances than low-cost housing. Perhaps HUD Secretary Romney had this in mind when he developed a program not with a number but with a name—Operation Breakthrough—and just recently announced the winners in the first round. It developed they came from thirty-seven states. They included the Aluminum Co. of America, U.S. Steel Corp., Westinghouse Electric, G.E., TRW. I would suggest that George Romney is in the process of doing what has to be done in an industrial society with a whole mass of countervailing powers. He is amassing a constituency within and without the business and the academic and the governmental communities for what he feels has to be done.

This brings me to my complete agreement with Professor Galbraith. We don't need further refinements of the analysis of facts when we're only 25 percent sure of the facts to begin with. I think there is something wrong with your profession when an honors graduate in economics who has followed the literature as best he could from time to time and is engaged in the pursuits you are studying picks up the *Quarterly Journal of Economics* and finds more figures than words and most of the words are explaining the figures. When you get done with the point of this overrefined analysis of uncertain data, it really doesn't matter. One of the God-awful dangers is that this infection is traveling to the business schools. A young fellow comes in and I talk to him about a job with our firm and I say, "What would you like to do?" And he says, "I'm very good at matrix algebra." Even in a great big company, how much matrix algebra can you use? And then systems analysis. As I understand systems analysis, everything is so interconnected with everything else you might as well do nothing. I have paid some attention to the field of housing and I discovered that at M.I.T. they have now applied systems analysis to the problems of the city. The answer they have reached is: don't build any low-cost housing because that really intensifies the problem. I would suggest, as Professor Galbraith has suggested, that if your conclusions come out ab-

surd, there is something the matter with your process. I would also suggest that you can't really predict the future. You're better off to take a plunge on a pilot project and then to monitor the results. Who in the department store business realized when he sensed that tremendous demand for maxicoats that first of all they would be worn with miniskirts? Did any sovereign producer of ladies clothing really sit back and say to himself that his ads could persuade women to walk around half naked but wrap themselves in blankets down to the ground? And if he did believe that, what systems analysts would have been far-sighted enough to warn, "When you start selling maxicoats, put on more house dicks because the shoplifting has become infinitely easier when you're walking around wrapped in a blanket."

What I am suggesting is that the economist should take himself away from the computer, away from the calculator, away from the vector analysis and all those things and get back to the real problems of our society. The responsible economist should do the kind of thing that Ken has suggested to you. I do disagree with him on some of his details, but I don't disagree with him as to the duty of today's economist. There are an awful lot of real problems in the world. There is that middle of the egg, literate and affluent. You, the educators, made them literate, and we, the industrial world, gave affluence. But at the bottom of the egg there is that uneasy, unhappy collection of the deprived who want in to the middle. We can't get them in unless we forge an effective political alliance between businessmen who don't oppose what's generally good for society although they may fight a small battle on some special issue that is particularly good for themselves. The essence of the matter is to bring economists into the problems of the day. As Professor Galbraith says: your purpose should be to affect belief and thereby to effect action. Or as Oliver Wendell Holmes, Jr., once put it: "Life is action and passion. I think it is required of a man that he should share the action and passion of his time at peril of being judged not to have lived."

HAROLD DEMSETZ: Professor Galbraith's paper begins with a discussion of the sociology of economics, continues with an attack on textbook economics, and concludes with a description of the view available from the vantage point of *The New Industrial State*. My comments are in agreement with the order of Galbraith's presentation only.

I am no expert on the sociology of economics and so I speak but briefly on this subject. Galbraith laments the resistance of the profession to new ideas and the persistence with which it holds to the old chestnuts. There may be much truth to his description of the sociology of the profession, although I am told that Ricardo's views were accepted rather quickly and that Mill's widely used text contained an early refutation of Say's law. But the frequency with which pretentious manuscripts prove to be incorrect or vacuous suggests that it is probably wise to adopt a conservative attitude toward our new ideas as well as those of others. An honest assessment of the promise and diffi-

culties of a radical transformation of economics or of society itself suggests that we should be cautious and humble.

Galbraith chides the profession for having held overlong to Say's law, which Galbraith describes not only as wrong but as having delayed progress in economics. The profession's persistence in this matter is much exceeded by the tenacity of Galbraith. The central proposition of *The New Industrial State*, which I quote from Galbraith's paper, is Say's law resurrected. Galbraith views the economic process as one "in which the ultimate accommodation in a significant part of the economy is to the producer. The individual's wants, though superficially they may seem to originate with him, are ultimately at the behest of the mechanism that supplies them. In the most obvious manifestation the producing firm controls its own prices in the market and goes beyond to persuade the consumer to the appropriate responding behavior." Supply creates its own demand in the views of both Galbraith and Say, but Say, alas, failed to explicitly include in his law the salesman.¹

I now come to a partial agreement with Galbraith, almost. Textbooks, even those brought up to date every two years, contain many errors, and their treatment of subject matter is ten years behind the profession. It is easier to criticize textbooks than to write them.

I agree that the concept of consumer sovereignty is misleading. However, Galbraith errs in identifying this concept as a correct deduction or assumption of modern economics. It is the property of textbook writers and of the same larger public upon whom Galbraith places such great reliance. It may well be that in the general view, as Galbraith puts it, "economics is a process by which the individual imposes his will on the producer." On occasion this view, as Galbraith further notes, may be offered by professional economists such as Fisher, Griliches, and Kaysen. No doubt many of us are guilty of such carelessness. Nonetheless, economics makes no such assumption.

What economics does assume is that exchange opportunities will be consummated when they appear advantageous to all parties to the exchange. Neither supplier nor demander, neither producer nor consumer, imposes his will on other parties to the exchange. The essence of market exchange is that all parties have only the right to offer for sale or offer to purchase. Exchange cannot be imposed on others.

The most elementary manipulation of supply and demand reveals that no exchange takes place, even when consumers desire the good, if they are unwilling to offer prices that cover the cost of supplying the good. Similarly, producers realize no sales if they are

unwilling to sell at prices as low as those being offered by consumers. Buyers who demand Conestoga wagons for \$5.00 a piece will be disappointed. So will suppliers who ask \$50,000.

Even in the long run, market survival demands no more from sellers than it does from buyers. Each can spend his way into bankruptcy and each can survive bankruptcy without charity only if he remains within his budget constraint. Neither buyer nor seller is sovereign in the economics of the market place.

Just as it is possible to deny consumer sovereignty without rejecting modern economics, so it is possible to deny the usefulness of perfect competition for some problems without thereby denying the relevance of rivalry. Galbraith is led to slight, if not to ignore, the very real fact of rivalry because modern industry does not seem to fit the assumptions of the perfectly competitive model. Partly this is the fault of the elementary texts that so often misuse and abuse the perfectly competitive model. It is surprising, nonetheless, that Galbraith gives virtually no attention to the problems created for his schema by rivalry among firms—even if all of them seek to maximize the panoply of organization interests, whatever that means. Each firm cannot grow as rapidly as it pleases. One industry must compete with another and so must firms within an industry. If the competition is not through price, then let it be through advertising or product development.

The effect of this rivalry is not so one dimensional as Galbraith would have us believe. There is no lack of appeal to leisure activities or to the arts. The airlines, the book and art clubs, the theaters, and the sports and hobby industries compete for our attention. This rivalry does not uniformly promote ever larger rates of consumption. Thrift and abstinence are encouraged by the banks, insurance companies, and, at least for the private expenditures of its citizens, by the sale of government bonds. Nor does this rivalry cater only to fashion change. Volkswagen and other compact car manufacturers actively promote infrequent changes in body style.

Because Galbraith fails to give this multidimensional rivalry the attention it deserves, his conclusion that our society is geared to overproduction of useless goods and underproduction of leisure must be suspect.

Suspicion is heightened when Galbraith argues seriously that "we have more appliances than houses because General Electric is more powerfully sovereign than the house builders." If this meaningless comparison is to be made, one would expect that attention might be called to the fact that the price of a house is about 1,000 times greater than the price of an appliance. A quick survey of my home, using Galbraith's method of reckoning producer sovereignty, revealed that the most powerful sovereign in American industry is the manufacturer of clothes hangers. Toothpick producers ran a close second.

But if neither consumer nor producer sovereignty is assumed in economics, it is true that in practice much of economic analysis does take wants as given. Mostly, this is the case because neither economics nor

¹ In *The New Industrial State* (Houghton Mifflin Co.), Galbraith devotes two complete chapters to a description of how supply creates its own demand. Chap. XVIII discusses the management of specific demand by the technostructure and Chap. XX, the regulation of aggregate demand.

behavioral science nor psychology nor yet Galbraith has developed a rich theory of wants. When a theory of want-formation is constructed, wants will no longer enter the analysis as exogenous variables. Even when such analysis can take place, economists, as scientists, will be in no position to question the "goodness" or "badness" of wants any more than today's chemist can question the morality of chemical affinities. This constraint is imposed by the scientific method, not by the subject matter to which it is applied.

To say that a theory of want-formation is lacking is not to say that economists take wants as "inborn" or "God-given." Madison Avenue, Washington, D.C., the Church, Mother, Galbraith, and now even Friedman, are but a few of the forces that form wants. There is surely little doubt that our beliefs and tastes are to a large extent learned, but it is not clear to me that business firms have an undue advantage in want-creating activity. Washington enjoys free press and TV coverage and is immune from antifraud laws. Professorial promoters are protected from the same laws by the camouflage of academic freedom.

Should the Playboy Club market a nicely packaged regimen that promises heaven to those who join, Mr. Hefner would soon find that his packages violate fair-packaging and labeling laws, if not indecent exposure laws, and that his promise amounts to fraud. Yet the churches, with impunity, offer to their members everlasting life if the correct regimen is followed. I resist the temptation to argue that these legal discriminations result in an imbalance in this nation's output, unduly favoring the public, religious, and educational sectors.

I now turn to the view offered by the concept of producer sovereignty. Galbraith resorts so often to undefined concepts and qualifying phrases that it becomes very difficult to isolate empirically meaningful hypotheses. For example, it never is really clear just how to rank large firms by the measure of producer sovereignty they wield. If Galbraith truly seeks the understanding of the profession, he should at least put his propositions into testable form. I have attempted to draw some inferences from his writing mainly by assuming that the subset of firms which possess the greatest producer sovereignty contains the large industrial firms. I trust that Galbraith will find that this assumption and its variants are reasonable interpretations of his work.

Galbraith claims that the part of the economy that is highly industrialized and highly organized, the part in which producer sovereignty is presumably greatest, will have as its organizational goal "peaceful relations with labor and the organizational security that comes with uninterrupted production." Such firms, Galbraith asserts, will be less resistant to wage demands which can easily be exported through higher prices "to the more vulnerable sector where consumer sovereignty is still relevant." It seems reasonable to infer from this that during an inflationary period prices will rise more rapidly than elsewhere in the industrialized, highly organized sector, where Galbraith claims the technostucture is strongest.

The implicit GNP price deflators, as of the third quarter 1969, show inflation adjustments of 142.5 for services, 142.0 for structures, 117.7 for goods output, and only 104.4 for gross auto product—just the reverse of what is suggested by producer sovereignty.

For September, 1969 (1957-59 = 100), the Consumer Price Index shows all items at 129.3, services at 146.0, shelter at 136.1, food at 127.5, and transportation at 123.6. Again it seems that the less industrialized sectors of the economy have produced the largest price increases. In the matter of price increases, a special theory of producer sovereignty seems needed to explain automobile prices which in September stood at only 99.5 percent of their 1957-59 levels.

If the industrialized, highly organized sectors of the economy are the stronghold of the technostucture and if, as Galbraith claims, the technostucture succeeds in persuading the public to overbuy its output, then we should expect to observe that these sectors of the economy are producing ever larger shares of national income. Assuming that the industrialized, highly organized industries are fairly represented by the manufacturing, transportation, communication, and utility industries we find that the share of national income produced in these sectors falls from 40.1 percent in 1950 to 38.7 percent in 1965. Including the federal government in the technostucture, we find that the share of national income produced in these sectors falls from 45.5 percent to 44.6 percent over the same time period.

According to Galbraith, the U.S. economy, because it overproduces goods, underproduces leisure. Yet, from 1890, the date that approximately marks for Galbraith the beginning of significant change in our industrial structure, the average number of hours worked weekly in U.S. manufacturing has declined by one-third, probably the most dramatic reduction in workload for large numbers of laborers that the world has experienced.

In *The New Industrial State* (p. 339), Galbraith writes: "Risks that would otherwise be unacceptable can be assumed in the civilian economy if they are protected by the much more nearly riskless weapons economy. . . . These advantages of the weapons competition to the industrial system could not easily be sacrificed by the industrial system." And in his present paper, Galbraith reaffirms the strength of producer sovereignty in the defense industries.

In preparation for this paper I analyzed a sample of defense stocks to see what evidence could be marshaled on this claim on short notice. The sample contains thirteen of the top prime defense contractors for whom prime defense contracts accounted for over 30 percent of sales.² These stocks over the period

² The stocks were taken from a list of top prime defense contractors found in *Fortune*, Aug. 1, 1969, p. 74. Two conditions for selection were set: (1) prime defense contracts must have accounted for over 30 percent of 1968 sales revenue; (2) the corporate history of the company must have been sufficiently uncomplicated that it could be traced quickly on the

1949-64 offered to investors about 21 percent more risk, measured by the mean deviation of the year-to-year rates of return, than did thirteen stock randomly selected portfolios.³

Galbraith's work, however, is done a disservice if it is viewed wholly from a scientific perspective. I do not think that Galbraith claims to have produced a carefully worked out scientific treatise.

But the extreme philosophical determinism implicit in his work should make the reader pause. Freedom

tapes of the Center for Research in Security Prices at the University of Chicago. The thirteen stocks are, in order of percentage of sales attributable to prime defense contracts, Lockheed, General Dynamic, Avco, Martin Marietta, United Aircraft, Grumman, Olin, McDonald-Douglas, Textron, Sperry-Rand, Honeywell, North American Aviation, and the Boeing Corporation. I would like to thank Professor Larry Fisher for putting aside his work to aid in this experiment.

³ The adjustments required to compensate for the lesser diversification of a thirteen stock portfolio relative to a market portfolio will be published by Lawrence Fisher and James H. Lorie in a forthcoming issue of the *Journal of Business*.

can have but little meaning in *The New Industrial State*. Captured by the technology of the modern affluent society, men, whether as consumers or voters, are persuaded of their wants by the technostucture. Galbraith would have us neglect these wants because, in some sense, they are not "natural" or "inborn."

But freedom surely means more than the right to exercise only elemental instincts. Freedom must mean the right to choose among the offerings of those who would persuade us. If choice is to be meaningful, a free society can rely only on a policy of not hindering the flow of communication from those who compete for the favor of grown men and who would persuade them of what is in their interest. And the free society must rely on competition to provide sufficient numbers of alternatives.

Peculiarly, the public at large, which Galbraith claims is the captive of the technostucture and which cannot be trusted to make choices in its self-interest, is the same public which he relies upon to recognize the truth of producer sovereignty; and this larger public, Galbraith predicts, will press this truth upon the educational and scientific estate—the very same estate to which Galbraith entrusts the control and guidance of the affluent society.

AMERICAN ECONOMIC ASSOCIATION

PROCEEDINGS
OF THE
EIGHTY-SECOND
ANNUAL
MEETING

NEW YORK, NEW YORK
DECEMBER 28-30, 1969

PROCEEDINGS OF THE
AMERICAN ECONOMIC ASSOCIATION

ANNUAL BUSINESS MEETING, DECEMBER 29, 1969
NEW YORK HILTON HOTEL, NEW YORK, NEW YORK

The Eighty-second Annual Business Meeting of the American Economic Association was called to order by President W. J. Fellner in the Trianon Ballroom of the New York Hilton Hotel, New York, New York, at 9:30 p.m. on December 29, 1969. At this point, a group of approximately twenty-five members of "a group of radical economists" filed into the meeting room. Their spokesman, Arthur MacEwan, demanded, on behalf of the group, that he be permitted to present a statement to the meeting. President Fellner informed Mr. MacEwan that he could present his statement when the meeting considered unfinished or new business. Even though ruled out of order by Martin Shubik, who announced that he had been asked to give advice as parliamentarian, and in the midst of confusion brought about by attempts of other members of the radical group to join him on the platform, Mr. MacEwan began immediately to read his statement. Meanwhile, President Fellner, using the microphone, proceeded to a consideration of the items as they appeared on the agenda prepared for the meeting. Despite the confusion occasioned by the fact that Mr. MacEwan was still reading his statement and was being interrupted by members of the audience, it was VOTED: (1) to approve the minutes of the Association's Business Meeting of December 29, 1968; (2) to ratify the minutes of the Executive Committee meetings held on December 27, 1968, and March 7-8, 1969; and (3) to accept the reports of the Secretary and the Treasurer, which had been circulated in written form to the members. (These reports are published below.)

At this point, President Fellner recognized Mr. MacEwan, who presented over the microphone the statement of the radical economist group, printed as Exhibit I to these minutes.

Following the reading of the statement, the group left the meeting. In the discussion that followed, the President recognized Gordon Tullock, who made the statement printed as Exhibit II to these minutes.

The President recognized Gerhard Tintner, who moved that a vote of censure of the

United States Department of State for refusing to grant a visa to the Marxian economist Mandel, who had been invited to make a series of lectures in the United States. The motion was seconded. Martin Shubik moved that the original motion should be amended to read, "It is the sense of this meeting that no member of the economics profession, regardless of his views, should be prevented from accepting an invitation to visit the United States to present lectures or to accept a teaching assignment." The motion was seconded. The question was raised as to whether the motion as amended was in violation of Article III, Section 3, of the charter of the A.E.A. that "the Association as such will take no partisan attitude, nor will it commit its members on practical economic questions." Following a discussion, it was voted 112 for and 7 against to adopt the motion as amended.

The Secretary presented the report of the Committee on Elections and the certification of the new officers for 1970 as follows:

In accordance with the bylaws on election procedure, I hereby certify the results of the recent balloting and present the reports of the Nominating Committee and the Committee on Elections.

The Nominating Committee, consisting of Fritz Machlup, Chairman, Carl F. Christ, Anne O. Krueger, W. Arthur Lewis, Robert V. Roosa, and Harold M. Somers, presented to the Secretary the list of nominees for the respective offices:

For President-elect

James Tobin

For Vice-Presidents

Evsey D. Domar
Homer Jones
Tibor Scitovsky
Edward S. Shaw

For Executive Committee

Robert Dorfman
Alain Enthoven
Margaret S. Gordon
Arnold C. Harberger

The Committee on Elections, consisting of Frank P. Brechling, Chairman, George E. Delehanty, and Harold F. Williamson, prepared biographical sketches of the candidates and ballots were distributed shortly after the middle of September. The canvass of the ballots was made on November 18, 1969, and the results were filed with the Secretary.

From the report of the Committee on Elections, I have the following information:

Number of envelopes without name for identification	135
Number of envelopes received too late ..	3

Number of defective ballots	—
Number of legal ballots	5,460

Number of returns from the mail ballot 5,598

On the basis of the canvass of the votes cast, I certify that the following persons have been duly elected to the respective offices:

President-elect (for a term of one year)

James Tobin

Vice-Presidents (for a term of one year)

Evsey D. Domar

Tibor Scitovsky

Members of the Executive Committee (for a term of three years)

Robert Dorfman

Arnold C. Harberger

Following the report on the elections, President Wassily Leontief took over the chair.

The President introduced President-elect James Tobin, who outlined his plans briefly for the 1970 annual meeting.

George Borts, Editor of the *American Economic Review*, summarized the operations of the Review Editorial Office during 1969. Mark Perlman, Editor of the *Journal of Economic Literature*, reviewed the operations of the Journal Editorial Office during 1969.

The President recognized Richard M. Bell, who asked that it be recorded that the dissidents who attended the Business Meeting represented a minority of the members of the Union for Radical Political Economics.

The President recognized Manuel Gottlieb, who asked whether there was any plan to pre-publish papers for future meetings. Following a discussion, it was agreed that the Executive Committee should consider the possibility of pre-publishing at least two complete sessions for the 1970 meeting.

The Secretary announced that the Executive Committee had VOTED to adopt the following schedule of membership dues, as of January 1, 1971: regular members, \$15.00; family members, \$1.00; junior members, \$7.50; contributing members, \$25.00 or more; and life members, \$300.00. The Executive Committee had also VOTED, effective immediately to accept no more separate subscriptions to the *Journal of Economic Literature*.

On behalf of the Executive Committee, the Secretary presented the following resolution, which was adopted:

This meeting would like to record a vote of special thanks to George McKinney, Chairman, Richard M. Cloney, James C. Cooper, Courtney M. Blackman, and other members of the Committee of Local Arrangements for their efficient management of the 1969 annual meeting of the Allied Social Science Associations.

The Secretary presented the following resolution, which was also adopted:

Be it resolved that this meeting commend Wassily Leontief and members of the Program Committee for planning such an excellent and varied program for the 1969 annual meeting.

The meeting was adjourned at 10:45 p.m.

EXHIBIT I

We have come to denounce the American Economic Association, and to denounce the dominant economics for which the A.E.A. provides the organizational support.

Economists in the United States work as a group and work contrary to the interests of the masses of people. The affluence and the power of the economists derive from their support of the elite, the elite which controls the institutional structure and the sources of power that perpetrate and reproduce the oppression of millions—the economists are the sycophants of inequality, alienation, destruction of environment, imperialism, racism, and the subjugation of women.

Economists are the priests and prophets of an unjust society. They preach the gospel of rational efficiency, justifying the reduction of man and nature to marketable commodities; they treat human beings as capital and tell us the poor are poor because they lack "productive skills"; all they tell us about the war in Vietnam is how to fight it more efficiently; they apply mathematical models that "prove" that foreign investment helps the development of poor countries; they tell us that racism is the result of "personal preference"; they tell us that private property and wage differentials present a system of personal material incentives "necessary" for "growth."

But the economists do not merely praise the system; they also supply the tools—indeed, they are the tools—instrumental to the elite's attainment of its unjust ends. They show how to manipulate people so that the system's hinges are smoothly oiled. Economists are minimizers of just discontent: in the face of police riots in cities, it is the economists who develop "people appeasement" programs to prevent rebellion; when a reactionary government controls a poor country, economists are sent to "rationalize" and "stabilize" its economy; when students rebel on campuses, it is the industrial relations economists and game theorists, the rational arm of the police, who provide the program for repression.

The American Economic Association must be denounced as the organization through which these economists operate. But further, the A.E.A. plays directly destructive roles in our society. It serves to insure the perpetuation of professionalism, elitism, and petty irrelevance. It serves to inhibit the development of new ideas, ideas which are reflective of social reality.

Our conflict with the A.E.A. is not simply an intellectual debate. The A.E.A. cannot lessen our condemnation by their willingness to partake in debate, or by

their willingness to provide a room to radical economists at this meeting. Our conflict is a basic conflict of interests. The economists have chosen to serve the *status quo*. We have chosen to fight it.

EXHIBIT II

May I begin by saying how happy I am that certain people are now leaving the meeting.

As those of you who know me are aware, the particular political and economic point of view which I happen to hold is a minority one, both in this society and in the larger society within which we live. I have become accustomed to being in the minority in majority voting. Nevertheless, I am here this evening to speak in favor of democracy within the American Economic Association—a system in which I almost always lose. Democracy may not be the best possible form of government either for a nation or for a voluntary association such as the A.E.A. You may recall that Winston Churchill once said democracy was the worst of all possible forms of government, except, of course, for those others that had been tried from time to time. Nevertheless, it is at the moment the best form of government which we have available. It is particularly suitable for voluntary societies such as this one, since minorities are free to leave and set up new societies if they find this one objectionable. It may be so that, over time, we will discover better ways of running such associations, but we have not done so at the moment.

Under the circumstances, it is necessary that we attempt to conduct our business in a reasonably orderly way. In democracy, everyone has a right to speak, but if they all speak at once, no one can be heard. Rules as to who speaks first are essential and cannot be objected to, provided they do not deprive anyone of the right to speak eventually. The disturbance which has just terminated did not arise because the disturbers were denied the right to speak, but simply because they were told they could not speak at exactly the time they chose without any consideration

for other persons' rights. They refused to obtain the floor by parliamentary methods apparently because they felt they had a superior right. This superior right, I presume, derives from the fortunate fact that they knew the truth and were here to proclaim it, whereas those who might otherwise wish to talk were in error. The rules which provide for the presenting of different points of view in an orderly manner would, thus, be unnecessary. The fact that the disrupters left immediately after reading their statement in order to prevent themselves from being contaminated by anything that anyone else might say is a further illustration of this attitude.

Clearly, the position taken by these people is profoundly undemocratic. They know the truth and propose to act on it, and they have no feeling that any particular procedure or routine has any value in itself. Any orderly process of making decisions will impose delay. It is certainly true that it would be better to reach the right decision instantaneously. Unfortunately, we have no instantaneous method of "aggregating preferences." Until one is invented, we will have to follow processes which can make errors and which take time.

As I said before, it is by no means obvious that democracy is the ultimate or ideal force of government. We may discover something better in the future. For the moment, however, we are stuck with democracy simply because it is better than any known alternative. It seems to me, therefore, very important that we not only conform to rules of democratic behavior ourselves, but enforce these rules on others. There is a certain irony in my pressing this point of view because, as I noted at the beginning of my remarks, I have almost always been in the minority in the past and it seems likely that I am doomed to be in the minority in the future. Nevertheless, until I can suggest some better method, I feel that I must accept the majority decisions and, as a matter of fact, the majority appears to be quite tolerant of various points of view, even the points of view of members of extremely small minorities like myself.

THE JOHN BATES CLARK AWARD

CITATION ON THE OCCASION OF THE
PRESENTATION OF THE MEDAL TO

MARC L. NERLOVE

DECEMBER 29, 1969

The John Bates Clark medal of the American Economic Association is awarded biennially "to that American economist under the age of forty who is adjudged to have made a significant contribution to economic thought and knowledge." For 1969 the medal is awarded to Marc Nerlove, whose life has been distinguished since childhood by the precocity of his achievements, including a baccalaureate with honors from the University of Chicago, at 18, a doctorate at 23 from Johns Hopkins, and professorships at Stanford, Yale, and now Chicago. In the highest traditions of economics as an empirical, quantitative, and cumulative science, his contributions have been creative and innovative both in method and in substance. His results are landmarks in the study of supply, demand, and production. His methods of estimating response lags and of analyzing time series have found general and fruitful application. With awe and gratitude for his record to date and with high expectations of his achievements when he grows up, I have great personal pleasure in presenting, on the Association's behalf, the Clark medal to Marc Nerlove.

MINUTES OF THE EXECUTIVE COMMITTEE MEETINGS

1. Minutes of the spring meeting held in New York City, March 7-8, 1969:

The first meeting of the 1969 Executive Committee was called to order at 9:30 a.m. in the New York Hilton Hotel, New York, New York, March 7, 1969. The following were present: W. J. Fellner, presiding, Moses Abramovitz, G. H. Borts, K. E. Boulding, Mary Jean Bowman, Otto Eckstein, Milton Friedman, Wassily Leontief, Mark Perlman, Anthony Scott, and H. F. Williamson. Absent were: K. J. Arrow, J. S. Bain, Franco Modigliani, and C. L. Schultze. Present as members of the Nominating Committee were: Fritz Machlup, Chairman, C. F. Christ, Anne O. Krueger, W. A. Lewis, R. V. Roosa, and H. M. Somers. Present as guests were: W. H. Nicholls and James Tobin.

1. *President's Remarks* (W. J. Fellner). President Fellner outlined the items on the agenda to be covered at the meetings.

2. *Minutes*. The minutes of the December 27, 1968, Executive Committee meeting were approved.

3. *Report of the Secretary* (H. F. Williamson). *Informational Ballot on Time and Place of Annual Meetings*. The results of the survey authorized by the Executive Committee showed that 3,498 voted to move the 1972 annual meeting from Chicago and 2,911 voted not to move. On the question of the time of the annual meeting, a plurality but not a majority of the voters indicated their preference for continuing to meet between Christmas and New Year's.

Following the discussion, it was VOTED to move the 1972 annual meeting from Chicago, and the Secretary was instructed to inquire specifically about the possibilities of Toronto, Montreal, or Denver. It was also VOTED not to change the time of the annual meetings.

The 1968 Annual Meeting in Chicago. The Secretary reported that the total registration for the annual meeting in Chicago was approximately 5,200. He reported the net income from the meeting was almost \$27,000, of which the share of the Association was a little over \$20,400. He also noted that the employment service was unusually busy. The total applicants registered was 2,061, employers 614, total vacancies listed 2,029, and approximate number of messages exchanged between employers and applicants 21,000.

Future Annual Meetings. The Secretary reported that plans were well under way for the 1969 annual meeting in New York City, and that hotel arrangements have been confirmed for the meetings in Detroit in 1970, New Orleans in 1971, New York City in 1973, and San Francisco in 1974.

Papers and Proceedings. It was announced that plans were to print 27,000 copies of the 1969 *Papers and Proceedings*. The total manuscript pages was 1,324.

The National Register of Scientific and Technical Personnel. It was announced that preliminary plans were under consideration for the participation of the Association in the 1970 National Register of Scientific and Technical Personnel.

Foreign Graduate Student Screening Project. It was announced that the Foreign Graduate Student Screening Project will be terminated as of the end of September, 1969.

4. *Report of the Treasurer* (H. F. Williamson). It was noted that the final auditor's report covering operations during fiscal 1968 was not yet available. A preliminary report, as of March 5, 1969, indicated that the expenses of the Association exceeded income by \$6,063 during 1968 and that the unrestricted balance as of December 31, 1968, was just over \$414,749.

5. *Report of the Editor, A.E.R.* (G. H. Borts). The Editor reported the completion of the transfer of editorial operations of the *American Economic Review* from Stanford University to Brown University. He raised two questions for consideration. One was whether the Association should continue to pay authors for articles accepted for publication in the *A.E.R.* The other was whether the *A.E.R.* should continue to exchange advertisements with other journals. Following a discussion, it was VOTED: (1) to discontinue payments to authors for all articles accepted after March 7, 1969, for publication in the *A.E.R.*; and (2) to discontinue all exchange advertisements as soon as feasible.

6. *Report of the Editor, J.E.L.* (Mark Perlman). The Editor reported that because of the problems involved in establishing his Editorial Office the March issue of the *Journal of Economic Literature* would not be published until April but that he expected all issues to be published on their regular schedule beginning in March, 1970. Because of the desirability of planning ahead, he asked for immediate approval of the 1970 budget of the *J.E.L.* rather than waiting until the December meeting of the Executive Committee. Following a discussion, it was VOTED to approve the following budget for the 1970 operation of the *J.E.L.*, subject to a review at the December, 1969, meeting by the Executive Committee of the overall financial needs of the Association for 1970:

Salaries	\$27,730*
3,000 hours of part-time personnel ..	9,000
Survey articles	3,000
Communications and supplies	2,200
	\$41,930

* These figures do not include fringe benefit costs.

Since further editions of the *Index of Economic Journals* are expected to be a by-product of the *J.E.L.*, it was VOTED to allocate \$4,578 to the *J.E.L.* to process articles that appeared in 1968 prior to the time that the *J.E.L.* began its collection.

At the request of the Editor, it was VOTED to appoint Marc Nerlove to the *J.E.L.* Board of Editors for a three-year term ending in 1972 to replace H. S. Houthakker, who resigned when he became a member of the Council of Economic Advisers.

7. Reports of Standing Committees.

Committee on Research and Publications (I. B. Kravis). In the absence of the Chairman, the Secretary reported on the difficulties encountered with the Pareto translation project. Following a discussion, it was agreed that the original plan to publish a translation of the *Manuel* should be dropped in favor of a volume of selected writings by Pareto.

Committee on Honors and Awards (James Tobin). The Chairman reported on the candidates recommended for the John Bates Clark Medal and the Distinguished Fellow award.

Nominating Committee (Fritz Machlup). The Executive Committee and the Nominating Committee met as an Electoral College to consider the nominees for the office of president-elect for 1970. After discussion, the nominee was selected and his acceptance obtained. Nominees for the other offices were discussed. The Chairman also submitted the recommendation of the Committee on Honors and Awards for the Distinguished Fellow award. The recipients were selected as the result of a ballot of the members of the Electoral College.

Committee on Economic Education (G. L. Bach). In the absence of the Chairman, the Secretary read a written report from G. L. Bach indicating that at its meeting in Chicago on December 28, 1968, the Committee had: (1) voted enthusiastically to authorize a resolution of appreciation for Ben Lewis' many years of service on the Committee; (2) strongly endorsed the plans for representatives of the A.E.A. to participate in the program on the problems of economic education in black colleges and universities to be held at Fisk University in April; (3) decided to recommend the postponement of the publication of the Joint Council on Economic Education first issue of the *Journal of Economic Education* until the fall of 1969 in order to make sure of the quality of its contents.

8. *Unfinished and Miscellaneous Business.*

a) President Fellner announced that Albert Rees had been renominated to serve as the Association's representative to the Social Science Research Council for a three-year term ending in 1972.

b) Upon recommendation of the Secretary, it was VOTED to reappoint S. J. Rubin for another three-year term as A.E.A. Counsel.

c) It was VOTED to dissolve the *ad hoc* Committee on Financial and Investment Policies and to extend a vote of thanks to the members for their services.

d) It was VOTED to instruct the Secretary to send letters on behalf of the Executive Committee to all retiring members of A.E.A. committees in recognition of their services.

e) President Fellner announced the appointment of K. E. Boulding as Chairman of the committee to handle the art exhibit at the 1969 annual meeting.

9. *Program for the 1969 Meeting* (Wassily Leontief). President-elect Leontief discussed his plans for the 1969 meeting, including the proposal to reproduce and distribute some of the papers prior to the meeting.

The meeting adjourned at 11:00 a.m. on March 8, 1969.

2. Minutes of the meeting held in New York City, December 27, 1969.

The second meeting of the 1969 Executive Committee was called to order at 2:00 p.m. in the New York Hilton Hotel, New York, New York, December 27, 1969. The following were present: William J. Fellner, presiding, K. J. Arrow, G. H. Borts, K. E. Boulding, Otto Eckstein, Wassily Leontief, Franco Modigliani, Mark Perlman, C. L. Schultze, Anthony Scott, and H. F. Williamson. Absent were: Moses Abramovitz, J. S. Bain, M. J. Bowman, and Milton Friedman. Present as guests were: J. S. Chipman, Robert Dorfman, A. C. Harberger, R. J. Lampman, J. P. Miller, W. F. Owen, Henry Rosovsky, Richard Ruggles, Tibor Scitovsky, G. J. Stigler, R. H. Strotz, James Tobin, and H. H. Villard.

1. *President's Remarks* (W. J. Fellner). President Fellner outlined the items on the agenda to be covered at the meeting. He also noted that an arrangement had been made for members of the Executive Committee to meet with the Black Caucus following the Richard T. Ely Lecture on December 28.¹

2. *Minutes*. The minutes of the March 7-8, 1969, meeting were approved. It was VOTED to hold the spring meeting of the Executive Committee in New York City, March 6-7, 1970, at the New York Hilton Hotel.

3. *Report of the Secretary* (H. F. Williamson). The Secretary announced that the Chairman of the Committee on Local Arrangements for the 1970 annual meeting to be held in Detroit December 28-30 would be appointed within a few weeks.

The schedule of future meetings of the Association was reviewed. At the recommendation of the Secretary, it was VOTED to plan to hold the 1972 meeting in Toronto, Canada. Following a discussion of the proposal to hold the 1975 meeting in Chicago, the Secretary was instructed to explore the possibility of holding the 1975 meeting in Washington, D.C., or Atlanta, Georgia, in order to avoid bad weather that might disrupt transportation to Chicago.

The Secretary reviewed other activities of the Association, which are treated more fully in the Secretary's Report printed below.

4. *Report of the Treasurer* (H. F. Williamson). The report of the Treasurer, covering the calendar year 1968 and published below, was summarized. Special attention was called to the fact that the excess of expenses over income for 1968 was \$6,063 and that the unappropriated surplus of the Association as of December 31, 1968, was \$414,749. The Treasurer pointed out that because of a rise in printing costs, the added expenses involved in the *Journal of Economic Literature*, and the unexpected costs incurred in connection with the 1969 *Handbook*, the excess of expenses over income during 1969 might well be close to \$83,000.

5. *Report of the Finance Committee* (C. W. Farnham). The report of this Committee, including a list of securities held by the Association as of December 31, 1969, is printed below.

6. *Report of the Editor of the American Economic Review* (G. H. Borts). The Editor reviewed the operations of the Editorial Office of the *American Economic Review*, covered in full below. It was VOTED to accept the report of the Editor including the nominations for the Editorial Board and the budget to cover the expenses of the Editorial Office for 1970.

7. *Report of the Editor of the Journal of Economic Literature* (Mark Perlman). The Editor reviewed the operations of the Editorial Office of the *Journal of Economic Literature*, covered in full below. It was VOTED to accept the report of the Editor, including the nominations for the Editorial Board and the budget to cover the expenses of the Editorial Office for 1970.

8. *Report of Standing and Special Committees.*

Committee on Research and Publications (J. P. Miller). The Chairman noted that he expected to call a meeting of the Committee in the near future to discuss operations. President Fellner summarized the status of the Pareto publication project. He recommended that the Committee be authorized to proceed

¹ A brief account of this meeting and a statement of concern prepared by the Black Caucus for publication in the *Papers and Proceedings* are printed below.

with plans to publish the volume of selected writings by Pareto. Following a discussion, it was VOTED to authorize the expenditure of \$7,000 (less any unspent balance from the original Ford grant) for the "Translation Series." John Chipman noted that the director of the Pareto fund was interested in the financing of the publication of a second volume of Pareto translations or a first-class translation of the *Manuel*.

Advisory Committee on the Index of Economic Essays (Richard Ruggles). The Chairman summarized the report on the economic index project printed below. He called special attention to the fact that a total amount of \$18,000 (in addition to the \$6,000 already authorized) would be needed to complete the projected volumes. Following a discussion, it was VOTED that after the money already appropriated has been spent, an additional amount of \$18,000 be allocated to be spent over a two-year period to complete the project.

Committee on Economic Education (G. L. Bach). The report of the Chairman is printed below. Also printed below is the report of the A.E.A. Visiting Scientists Program, prepared by Phillip Saunders.

Advisory Committee on the History of the Association (G. J. Stigler). On behalf of the Committee, the Secretary reported that no date had yet been set for the completion of the history of the Association.

Joint Committee with the Association of American Law Schools (G. M. Meier). In the absence of the Chairman, the Secretary summarized the report printed below. He noted the request that the Committee be authorized, in cooperation with the representative of the Association of American Law Schools, to proceed with efforts to obtain financial support for the activities of the joint committee. Following a discussion, it was VOTED to authorize this request.

Ad Hoc Committee to Review the Functions of Annual Meetings (K. J. Arrow). The Chairman reviewed the report printed below. Following a discussion, it was decided not to act at this time on the recommendations made by the Committee.

Economics Institute Advisory and Policy Board (Henry Rosovsky). The Chairman reviewed the report printed below. He recommended that in view of the possibility of the formation of an Institute for Black Students, modeled after the Economics Institute, the President appoint two representatives of the Executive Committee to serve on the advisory committee of such an institute. Following a discussion, President Fellner appointed Charles L. Schultze and Arnold C. Harberger to serve.

Committee to Nominate Candidates for the Office of Secretary-Treasurer (G. J. Stigler). The Chairman reviewed the work of the Committee and listed the names of the candidates recommended for consideration by the Executive Committee. Following a discussion, it was VOTED to authorize the following committee to approach the candidates and to offer the appointment: G. J. Stigler, Chairman, W. J. Fellner, and Wassily Leontief. It was further VOTED to thank the members of the original committee for their work.

9. Reports of A.E.A. Representatives and Advisory Committees.

A.C.L.S. (R. L. Andreano). No report.

S.S.R.C. (Albert Rees) Report printed below.

N.B.E.R. (W. L. Thorp). Report printed below.

A.A.A.S. (Carl Kaysen). No report.

I.E.A. (Fritz Machlup). Report printed below.

Census Advisory Committee (B. G. Hickman). No

report.

National Research Council, Division of Behavioral Sciences (J. N. Morgan). No report.

Consortium of Professional Associations for the Study of Teacher Improvement Programs (H. H. Villard). Henry Villard outlined the activities of CONPASS over the past year as they involved teachers and the teaching of economics. He also outlined the recommendations of the Committee of Economists that participated in the Institute sponsored by CONPASS at the Grove Park Inn, Asheville, North Carolina, June 10-15, 1969 (printed below).

10. Unfinished and Miscellaneous Business.

The Secretary raised a question as to whether an attempt should be made to prepublish papers to be given at the 1970 annual meeting. It was recommended that the President-elect explore this possibility in consultation with representatives of the other allied social science associations.

Wassily Leontief raised the question regarding the addition of the session on black economists to the sessions already scheduled to be published in the *Papers and Proceedings* of the 1969 annual meeting. Following a discussion, it was VOTED to authorize Professor Leontief to decide on whether this session should be included.

The Secretary read a letter from P. L. Ward, Executive Secretary of the American Historical Association, asking if the A.E.A. would be willing to contribute to the cost of employing a qualified individual to present the position of the learned societies in respect to the proposed taxation of foundations at congressional hearings on the tax reform bill. Following a discussion, it was VOTED to authorize the payment of \$100 if requested by the committee under the chairmanship of Paul Ward.

The Secretary read a letter from Wallace F. Lovejoy, Southern Methodist University, asking if the Association would be willing to contribute to a fund for the widow of Paul Homan, one-time editor of the *American Economic Review*. Following a discussion, it was decided that it was not appropriate to make such a contribution.

The Secretary read a request from J. W. Lehman, Executive Director of the American Statistical Association, asking the American Economic Association to join the A.S.A. and allied groups in sponsoring an invitation to the International Statistical Institute to meet in the United States in 1971 with the understanding that there would be no financial obligation involved. It was VOTED to join in cosponsoring this meeting, as requested.

The Secretary raised the question of granting complimentary memberships to editors of the *American Economic Review* and the *Journal of Economic Literature* upon retirement. Following a discussion, it was VOTED to adopt this policy and to give complimentary memberships to former editors J. G. Gurley and Arthur Smithies.

The Secretary read a letter from Daniel Thorner, indicating that the widow of A. V. Chayanov was living in Russia on a very small income. In view of the fact that publication of Chayanov's book on the peasant economy was published in the A.E.A. "Translation Series" without a payment for royalties, Thorner suggested that the A.E.A. might wish to make a small contribution to Madame Chayanov. Following a discussion, it was VOTED to make a contribution of \$300 to Madame Chayanov.

The meeting was adjourned at 12:00 m.

REPORT OF THE SECRETARY FOR THE YEAR ENDING DECEMBER 31, 1969

Annual Meetings. The final report for the 1968 annual meeting of the Allied Social Science Associations, held in Chicago, Illinois, indicated a total net registration of 5,136 individuals, of which 3,639, or 71 percent, preregistered. The net income from the meeting was \$27,547.26. The Association's share, based on registration figures, was \$20,443.81.

The schedule for future annual meetings is: 1970, Detroit, Sheraton-Cadillac; 1971, New Orleans, Jung; 1972, Toronto, Royal York (tentative); 1973, New York, New York Hilton; 1974, San Francisco, San Francisco Hilton; 1975, undecided; 1976, Philadelphia, tentative.

Membership. Exhibit I below shows that the total number of members and subscribers as of December 31, 1969, was 26,086, a net increase of 1,373 for the year.

EXHIBIT I
MEMBERS AND SUBSCRIBERS

	Totals 12/31/68	Gain or Loss	Totals 12/31/69
Class of membership:			
Annual.....	15,085	951	16,036
Junior.....	2,071	237	2,308
Family.....	248	26	274
Complimentary...	146	5	151*
Life.....	268	8	276
Honorary.....	17	1	16
Total members..	17,835	1,226	19,061
Subscribers.....	6,878	147	7,025
Totals.....	24,713	1,373	26,086

* Includes 89 who do not receive the publications.

Handbook. Publication of the 1969 edition is scheduled for January, 1970. The format is essentially the same as the 1964 edition. Main features are biographical sketches of all members who returned the questionnaires, classification of respondents by fields of specialization, classification by academic affiliation, classification of members by geographical location, and selected list of departmental chairmen. The total number of questionnaires re-

turned for the 1969 edition was 12,744, compared with 8,380 for the 1964 edition.

National Register of Scientific and Technical Personnel. Present plan is to mail questionnaires in February, 1970, to approximately 22,000 individuals who seem to be qualified as economists. Total includes all American Economic Association members, plus nonduplicates from membership lists of the American Agricultural Economic Association, the Economic History Association, and the Midwest, Southern, and Western Economic Associations. It is hoped that all members will respond promptly to the 1970 questionnaires. The information provided by the Survey is a most valuable source of the data needed to analyze the structure of the profession and how it has changed over the past several years.

Foreign Students Screening Project. During 1969, work was completed on the Second Edition of the book on *Graduate Study in Economics*. It includes descriptions of the graduate school offerings of 150 departments of economics and agricultural economics in the U.S. and Canada. An arrangement has been made with Richard D. Irwin, Inc., to distribute this volume.

With the termination of the Ford Foundation grant, the Association discontinued the operation of the Foreign Student Screening Project as of September, 1969. The possibility of continuing the screening operations under the Economics Institute is being explored.

Employment Service. The Association continues to sponsor three types of employment services. One is the "Vacancies and Applications" section which appears in each issue of the *American Economic Review*. The second is the National Registry for Economists, operated on a year-round basis by the Chicago Placement Service of the Illinois State Employment Service in Chicago. The third is the employment service provided by the state offices of the U.S. Employment Service at our annual meetings. Data on the operation of the employment service at the Chicago meeting indicated that the number of job openings

filed was 2,029; the number of registered applicants was 2,061. Nearly 21,000 messages were handled and 1,825 interviews were arranged.

Group Flights. Arrangements were made to provide group flights for members attending the 1969 meeting in New York City. It is expected that this arrangement will be continued for future meetings.

Use of the Mailing List. The continued policy of the Association is to grant requests for the use of the mailing list only if the material to be distributed appears to be of interest to a substantial proportion of our members. Net income from the sale of the mailing list during 1969 was \$12,591 compared to \$12,008 for 1968.

Advertising and Announcements. There were 179 pages of paid advertising and 13½ pages of exchange advertising (the exchange advertisements were discontinued after the March, 1969, issue of the *A.E.R.*) in the *American Economic Review* and the *Papers and Proceedings* compared to 184½ paid and 54 exchange pages in 1968. Advertising in-

come was \$50,272 compared to \$48,794 in 1968.

Permission to Reprint and Translate. Official permission to quote from, reprint, or translate and reprint articles from the *American Economic Review* and the *Papers and Proceedings* totaled 183 in 1969 as compared to 155 in 1968. Some 42 permissions granted in 1969 were in response to requests to reproduce or reprint articles for classroom or conference use. Attention should be called to the fact that upon receipt or a request to reprint, the publisher-editor is instructed to get the author's permission and then to send a copy of this letter of consent to the Association for official permission. It is also suggested that a fee of \$100 be paid to the author. While the Secretary's Office has no record of payments made, the correspondence received indicated that a considerable number of the authors have received this fee.

Papers and Proceedings. A comparison of the size and cost of the *Papers and Proceedings* for the period 1957-69 is shown in Exhibit II.

EXHIBIT II
PUBLICATION COSTS

Papers and Proceedings				Handbooks		
Year*	Number of pages	Number of copies	Cost	Number of pages	Number of copies	Cost
1957	754	12,400	\$16,253	548	10,100	\$15,815
1958	677	12,700	15,471	32	9,300	1,434
1959	689	14,000	16,780			
1960	745	14,800	18,914			
1961	675	15,800	19,759			
1962	615	16,000	18,277			
1963	753	17,700	23,440			
1964	698	18,500	23,362	472	16,000	48,626
1965	652	20,000	23,264			
1966	670	22,500	28,405	184	23,000	13,261
1967	741	24,000	32,682			
1968	751	26,500	36,637			
1969	629	27,000	41,263			

* This is the year of publication and pertains to the meeting of the preceding year.

*Standing Committees*ADVISORY COMMITTEE ON THE HISTORY OF
THE ASSOCIATION

George J. Stigler, *Chairman*
Joseph Dorfman
Harold F. Williamson

CENSUS ADVISORY COMMITTEE

Bert G. Hickman, *Chairman* (1972)
Morris A. Adelman (1969)
Daniel Creamer (1972)
Donald J. Daly (1971)
Solomon Fabricant (1972)
Gary Fromm (1970)
Raymond W. Goldsmith (1972)
Fred H. Klopstock (1969)
Sherman J. Maisel (1971)
Robert R. Nathan (1969)
Douglass C. North (1969)
Arthur M. Okun (1972)
Guy H. Orcutt (1970)
Murray L. Weidenbaum (1971)
Ernest W. Williams (1969)

CENSUS ADVISORY COMMITTEE, TECHNICAL
SUBCOMMITTEE ON BUSINESS CYCLE
DEVELOPMENTS

Geoffrey H. Moore, *Chairman* (1972)
Donald J. Daly (1971)
Gary Fromm (1971)
Gottfried Haberler (1970)
Bert G. Hickman (1970)
Frank E. Morris (1969)
Arthur M. Okun (1972)
Beryl W. Sprinkel (1971)
Harry Stark (1969)
Kenneth B. Williams (1971)

COMMITTEE ON ECONOMIC EDUCATION

G. L. Bach, *Chairman* (1970)
Kenneth E. Boulding (1969)
Marshall Colberg (1969)
Rendigs Fels (1969)
Allen C. Kelley (1969)

Henry H. Villard (1970)
Harold F. Williamson, *Ex Officio*

COMMITTEE ON HONORS AND AWARDS

James Tobin, *Chairman* (1970)
Richard A. Musgrave (1970)
William H. Nicholls (1972)
Douglass C. North (1972)
Anthony D. Scott (1970)
George P. Schultz (1972)

COMMITTEE ON RESEARCH AND PUBLICATIONS

John Perry Miller, *Chairman* (1971)
Zvi Griliches (1970)
Bert F. Hoselitz (1971)
James N. Morgan (1971)
Albert Rees (1969)
Willard L. Thorp (1970)
Harold F. Williamson, *Ex Officio*

ECONOMICS INSTITUTE, POLICY AND ADVISORY
BOARD

Henry Rosovsky, *Chairman* (1971)
Douglass C. North (1970)
Gustav Ranis (1972)
Vernon W. Ruttan (1970)
Joseph J. Spengler (1971)
Jeffrey G. Williamson (1972)
Wyn F. Owen, *Ex Officio*

JOINT COMMITTEE WITH THE ASSOCIATION OF
AMERICAN LAW SCHOOLS

Gerald M. Meier, *Chairman*
Louis De Alessi
Harold Demsetz
Richard A. Musgrave
Peter O. Steiner

JOINT COUNCIL ON ECONOMIC EDUCATION,
SPECIAL ADVISORY COMMITTEE

G. L. Bach (1970)
Rendigs Fels (1971)
Henry H. Villard (1972)

Committees Appointed During the Year

COMMITTEE ON ELECTIONS

Frank P. Brechling, *Chairman*
George E. Delehanty

Harold F. Williamson

AD HOC COMMITTEE TO SUGGEST NOMINEES
FOR SECRETARY-TREASURER

George J. Stigler, *Chairman*
Robert J. Lampman
Robert H. Strotz

FINANCE COMMITTEE

C. Wells Farnham, *Chairman*
Corliss D. Anderson
Milton Friedman
Harold F. Williamson

NOMINATING COMMITTEE

Fritz Machlup, *Chairman*
Carl F. Christ
Anne O. Krueger
W. Arthur Lewis
Robert V. Roosa
Harold M. Somers

Council and Other Representatives

A.A.A.S.

Joseph A. Pechman (1971)

A.C.L.S.

Ralph L. Andreano (1970)

COMMITTEE ON QUANTITATIVE DATA

Richard A. Easterlin

CONSORTIUM OF PROFESSIONAL ASSOCIATIONS
FOR THE STUDY OF TEACHER IMPROVE-
MENT PROGRAMS

Henry H. Villard (1972)
Harold F. Williamson (1970)

I.E.A.

John T. Dunlop (1973)
Fritz Machlup (1970)

NATIONAL ARCHIVES ADVISORY COUNCIL,
GENERAL SERVICES ADMINISTRATION

Harold F. Williamson (1969)

N.B.E.R.

Willard L. Thorp (1970)

NATIONAL COMMITTEE FOR THE REVISION OF
HANDBOOK II

Ernest Bartell (1969)
Henry M. Levin (1969)

NATIONAL RESEARCH COUNCIL, NATIONAL
ACADEMY OF SCIENCES

James N. Morgan

S.S.R.C.

Zvi Griliches (1970)
James N. Morgan (1971)
Albert Rees (1969)

*Representatives of the Association on Various Occasions*NATIONAL CONFERENCE ON TEACHER EDU-
CATION AND PROFESSIONAL STANDARDS

John E. Maher

SOUTHERN ILLINOIS UNIVERSITY, CHARTER
DAY CONVOCATION

Howard S. Dye

AMERICAN ACADEMY OF POLITICAL AND SO-
CIAL SCIENCES

Wilfred Malenbaum

UNIVERSITY OF VIRGINIA, SESQUICENTENNIAL
CONVOCATION

Thomas S. Berry

INAUGURATIONS

Laurence C. Smith, Westmar College
Kenneth L. Bauge
M. G. Scarlett, Middle Tennessee State Uni-
versity
Rudolph C. Blitz
G. L. Taggart, Utah State University
Owen Clark
Vivan W. Henderson, Clark College
James Crawford
Kermit A. Johnson, Alabama College
James F. Dostor
Ronald G. Weber, Mt. Union College
E. K. Eberhart

Lawrence Park, Mansfield State College
 Grant N. Farr
 Eric N. Rackham, McKendree College
 Kanji Haitani
 William H. Duncan, Millersville State College
 Khal M. Hamid
 Alfred O. Canon, Drury College
 Robert S. Hardie
 Gordon C. Bjork, Linfield College
 Lafayette G. Harter, Jr.
 E. Milton Grassell, College of Great Falls
 Robert I. Hislop
 Norbert J. Hruby, Aquinas College
 Frank H. Jackson
 John J. Pruis, Ball State University
 Charles Leese
 Herbert Schueler, Richmond College
 Jacob Loft
 William G. Caples, Kenyon College
 C. A. Monteith
 Robert S. Eckley, Illinois Wesleyan University
 Douglas Poe
 F. Palmer Sample, Lebanon Valley College
 Robert C. Riley
 Marion D. Thorpe, Elizabeth City College
 Stephen P. Shao
 Martin L. Shotzberger, Catawba College
 David H. Shelton
 Victor G. Rosenblum and Gregory B. Wolfe,
 Reed College
 Paul B. Simpson
 John O. Eidson, Georgia Southern College
 John L. Snare
 John T. Bernhard, Western Illinois University
 Douglas Y. Thorson
 William James McGill, University of California, San Diego
 Marjorie S. Turner
 Richard C. Jones, State University College at
 Cortland
 John E. La Tourette

W. Stewart Tosh, Mohawk Valley Community
 College
 Virgil Crisafulli
 Thomas A. Bartlett, Colgate University
 Melvin A. Eggers
 Roy Joe Stuckey, Jamestown College
 David C. Nelson
 Harold Brown, California Institute of Technology
 Harold M. Somers
 David Bicknell Truman, Mount Holyoke College
 Williard L. Thorp
 Robert John Henle, Georgetown University
 John W. Lehman
 John W. Gustad, Fort Hayes Kansas State
 College
 Daniel G. Rupp
 Joel Pritchard Smith, Denison University
 William Henderson
 David Loren McKenna, Seattle Pacific College
 Bayard O. Wheeler
 Gus Turbeville, Coker College
 Charles E. Edwards
 Frank R. Barth, Gustavus Adolphus College
 Morgan I. Thomas
 Richard L. Bowen, University of South Dakota
 Sammie Hasegawa
 Robert J. Giroux, Clarke College
 Peter Brust
 Charles Vevier, Adelphi University
 Ronan Macdonald
 Arthur Gene Hansen, Georgia Institute of Technology
 Edwin Frank Ulveling
 Jerald C. Walker, John J. Pershing College
 Wallace C. Peterson

HAROLD F. WILLIAMSON, *Secretary*

REPORT OF THE TREASURER FOR THE YEAR ENDING DECEMBER 31, 1969

The detailed results of the financial operation of the Association for the twelve-month period ending December 31, 1969, are shown in the Auditor's Report, printed below. As indicated in Exhibits 3 and 4, expenses exceeded income for 1968 by \$6,063, and as shown in Exhibit 1, the unappropriated surplus of the Association at the end of 1968 was \$414,749. A summary comparison of the major income and expense items shown in Table 1 reveals the extent to which a sharp rise in printing

costs, along with higher editorial and administrative costs, accounted for the shift from an excess of income over expenses in 1967 to a deficit in 1968.

Tables 2 and 3 provide summaries of the investment portfolios and return on investments for the years 1945-68. A list of the security holdings of the Association for 1969 is shown in the Finance Committee Report, presented below.

HAROLD F. WILLIAMSON, *Treasurer*

TABLE 1
SUMMARY OF ASSOCIATION'S INCOME AND EXPENDITURES, 1967-68

	1967	1968
Income		
Dues and activities	\$308,284	\$320,494
Investment income	39,327	44,623
Total	\$347,611	\$365,117
Expenses		
Publications		
Printing and mailing		
<i>American Economic Review</i>	\$ 81,984	\$112,803
<i>Journal of Economic Abstracts</i>	51,220	53,194
<i>Papers and Proceedings</i>	32,682	36,637
Editorial Offices: Salaries and expenditures ..	49,095	61,512
Other	10,000	8,400
Total publication expenses	\$224,981	\$272,552
Administration	88,408	98,628
Total	313,389	371,180
Excess of income over expense	\$ 34,222	\$ (6,063)

(Tables 2 and 3 will be found on page 500)

AMERICAN ECONOMIC ASSOCIATION

TABLE 2
INVESTMENT PORTFOLIO

Year	At Par	Cost			Market
	Bonds	Bonds	Stocks	Total	Stocks and Bonds
1945	\$ 40,000	\$ 36,705	\$ 44,955	\$ 81,661	\$103,574
1948	35,000	33,108	48,624	81,732	84,841
1950	35,000	33,108	51,978	85,087	104,177
1951	43,000	43,340	49,764	93,104	117,316
1952	42,000	42,312	58,934	101,246	130,836
1953	68,000	68,308	46,458	114,766	134,562
1954	61,000	61,518	38,082	99,600	132,280
1955	75,000	75,370	59,394	134,764	166,772
1956	75,000	75,370	60,237	135,607	168,337
1957	75,000	75,370	55,084	130,454	151,638
1958	75,000	75,370	67,741	143,111	175,609
1959	75,000	75,386	67,652	143,038	191,506
1959*	175,000	175,616	67,652	243,268	291,506
1960*	160,000	160,508	94,910	255,418	299,768
1961*	170,000	169,794	109,071	278,865	356,131
1962*	125,000	125,367	116,699	242,066	293,039
1963*	90,000	90,367	119,524	209,891	284,160
1964*	150,000	149,198	131,249	280,447	371,556
1965*	125,000	124,048	157,836	281,884	409,653
1966*	143,000	142,138	241,075	383,213	491,324
1967	110,000	109,799	249,672	359,471	546,024
1968	110,000	109,800	289,098	398,898	638,763

* Includes bonds held in temporary operating fund.

TABLE 3
RETURN ON INVESTMENTS

Year	Bonds	Stocks	Total	Rate of Return on Cost
1945	\$1,479	\$2,488	\$3,968	4.71%
1948	1,194	2,944	4,139	5.06
1950	1,117	3,860	4,977	5.85
1951	1,026	4,607	5,633	6.05
1952	1,117	3,681	4,799	4.75
1953	1,435	3,587	5,022	4.36
1954	1,621	2,961	4,582	4.58
1955	1,750	3,002	4,752	3.53
1956	1,770	3,336	5,106	3.76
1957	1,770	3,397	5,167	3.90
1958	1,770	3,182	4,952	3.46
1959*	2,518	3,231	5,749	3.90
1959†	3,894	3,231	7,125	2.90
1960†	6,693	3,772	10,465	4.09
1961	5,460	4,143	9,603	3.44
1962	4,838	4,489	9,327	3.85
1963‡	3,320	5,041	8,361	3.98
1964‡	3,341	6,364	9,705	3.46
1965‡	5,286	6,330	11,616	4.12
1966‡	6,356	7,957	14,312	3.73
1967	7,906	11,434	19,340	5.38
1968	8,444	11,231	19,675	4.95

* Does not include income from bonds held in temporary operating fund.

† Includes income from bonds held in temporary operating fund.

‡ Does not include interest on savings account.

REPORT OF THE FINANCE COMMITTEE

*Executive Committee
American Economic Association
Evanston, Illinois*

Gentlemen:

The accompanying inventory lists the securities held by the American Economic Association as of December 31, 1969, with the costs and the market values as of December 31. The total market value of the securities account as of December 13, 1968, was \$638,763. Value as of December 31, 1969, was \$614,140.

However, the value as of year end includes a \$170,000 Special Addition that was made in January of 1969 as part of a grant from the Ford Foundation which was commingled with the Association's account. It also reflects a total of \$145,000 which was withdrawn from the regular account during 1969. When these factors are taken into account, the value of the Association's portfolio as of December 31, 1969, was \$449,853 or 73.2 percent of the aggregate. The value of the Special Addition as of December 31, 1969, was \$164,287 or 26.8 percent of the aggregate.

After adjusting for the additions and withdrawals, we estimate that the market value of the aggregate account decreased by 6.88 percent during 1969. As a basis of comparison, the Dow-Jones Industrial Average decreased by 15.2 percent and the Standard & Poor's average of 425 stocks decreased by 10.2 percent.

During the current year, the bond portion of the investment account was increased in order to participate in the historically high yields which prevailed and to anticipate the prospect of lower interest rates in 1970. The following bonds were purchased: 20M Avco

7.50 11/1/93 and 20M Marcor 6.50 10/1/88. In terms of the common stock portion of the investment account, a reduction was made during the course of the year as it became increasingly clear that the Federal Reserve Board did not intend to let up on its restrictive monetary policies until there was clear evidence that the forces of inflation had been brought under control.¹ Among issues sold were: 400 St. Paul Companies, 200 Eastern Air Lines, 150 A. G. Bayer, 110 A. G. Siemens, 200 Swedish Ball Bearing, 275 Peoples Gas, 200 Standard Oil of Indiana, 200 Union Oil Convertible Preferred, 150 Marcor, 150 Inland Container, 300 United Air Lines, 200 Marathon Oil and 200 Dow Chemical. Among issues purchased were: 15M McDonnell Douglas Convertible Debentures, 20M National Equipment Rental Convertible Debentures, 15M United Utilities Convertible Debentures, 200 J. Ray McDermott, 300 Sherwin Williams, 200 Pennzoil United Convertible Preferred, 150 Honeywell, 100 Control Data, 250 SRF Fund, 100 Houston Lighting & Power, 300 Archer Daniels Midland and 400 American Hospital Supply.

At this time, while the Finance Committee's decision to maintain a large equity exposure seems warranted, the undersigned also believes that long-term bonds represent an extremely attractive investment and thus are appropriate in the Association's account, even though the controlling long-run investment policy is to be aggressive.

C. WELLS FARNHAM, *Chairman*

¹ This reduction is not apparent from the accompanying inventory because \$110,000 was withdrawn from the account's cash reserve during November and December.

AMERICAN ECONOMIC ASSOCIATION

INVENTORY AND APPRAISAL OF SECURITIES AND CASH AS OF 12/31/69

	Par or Shares	Market Value	Cost
CASH AND SHORT-TERM SECURITIES			
CASH			
<i>Corporate and Other</i>			
General Finance Notes.....	9,000	9,000	9,000
TOTAL CASH AND SHORT-TERM SECURITIES.....		9,000	9,000
LONG-TERM BONDS AND PREFERRED STOCKS			
<i>Corporate and Other</i>			
Anheuser Busch.....	10,000	8,100	10,000
Avco.....	20,000	15,600	17,850
Corn Products.....	10,000	7,600	9,950
General Tel. and Elec.....	20,000	15,400	20,530
Marcor, Inc.....	20,000	15,400	16,975
Pet, Inc.....	10,000	7,700	9,900
TOTAL LONG-TERM BONDS AND PREFERRED STOCKS.....		69,800	85,205
TOTAL CASH AND FIXED INCOME SECURITIES.....		78,800	94,205
LIMITED RISK CONVERTIBLES AND UTILITY			
COMMON STOCKS			
<i>Convertible Bonds</i>			
United Utilities.....	15,000	13,650	17,438
<i>Convertible Preferred Stocks</i>			
General American Transportation.....	400	19,200	13,862
<i>Utility Common Stocks</i>			
Central & South West.....	300	12,000	2,100
Houston Lighting and Power.....	300	11,700	4,992
Public Service Indiana.....	500	18,500	24,215
		42,220	31,307
TOTAL LIMITED RISK CONVERTIBLES AND UTILITY COM-		75,050	62,607
MON STOCKS.....			
OTHER COMMON STOCKS AND CONVERTIBLES			
<i>Financial</i>			
First Bank System.....	400	22,400	12,690
Wells Fargo.....	300	14,700	10,241
		37,100	22,931
<i>Food and Containers</i>			
Archer Daniels Midland.....	300	16,800	16,909
<i>Machinery and Construction</i>			
Deere and Co.....	300	12,900	9,484
McDermott, J. Ray, Co.....	200	10,600	17,395
McDonnell Douglas.....	15,000	10,950	17,775
Rex Chainbelt, Inc.....	300	9,300	6,620
Sherwin Williams.....	300	14,100	16,771
		57,850	68,045
<i>Transportation</i>			
National Equipment Rental.....	20,000	10,400	18,400
<i>Mining and Metals</i>			
International Nickel.....	250	10,750	3,911
McIntyre Porcupine Mines.....	• 100	14,400	4,818
		25,150	8,729
<i>Oil and Gas</i>			
Atlantic Richfield Co.....	200	17,200	9,625

REPORT OF THE FINANCE COMMITTEE

503

TABLE 1 (Continued)

	Par or Shares	Market Value	Cost
Pennzoil United.....	200	9,600	15,293
Royal Dutch.....	550	21,450	15,955
		48,250	40,873
<i>Chemicals and Drugs</i>			
Abbott Laboratories.....	300	22,800	6,135
<i>Electrical Products</i>			
Honeywell, Inc.....	150	21,150	17,562
Motorola, Inc.....	150	20,250	6,745
Westinghouse Electric.....	200	11,600	11,615
Zenith Radio.....	300	9,900	1,975
		62,900	37,897
<i>Office Equipment</i>			
Control Data.....	100	11,600	14,141
International Business Machines.....	122	44,408	10,169
Xerox Corp.....	150	15,900	5,305
		71,908	29,615
<i>Miscellaneous</i>			
American Hospital Supply.....	400	18,000	15,754
Disney Walt Productions.....	473	63,382	8,429
R. R. Donnelley and Sons.....	450	11,250	9,795
SRF Fund, Inc.....	250	14,500	17,852
		107,132	51,830
TOTAL OTHER COMMON STOCKS AND CONV.....		460,290	301,364
TOTAL SECURITIES AND CASH.....		614,140	458,176

REPORT OF THE AUDITOR

To the Executive Committee of American Economic Association:

We have examined the statement of assets and liabilities of American Economic Association (a District of Columbia corporation, organized not for profit) as of December 31, 1968, and the related statements of changes in restricted fund balances and income and expenses for the year then ended. Our examination was made in accordance with generally accepted auditing standards, and accordingly included such tests of the accounting records and such other auditing procedures as we considered necessary in the circumstances.

In our opinion, the accompanying statements present fairly the assets and liabilities of American Economic Association as of December 31, 1968, and the changes in restricted fund balances and income and expenses for the year then ended, in conformity with generally accepted accounting principles applied, except for the change (which we approve) in accounting for investments as described in the note to the financial statements, on a basis consistent with that of the preceding period.

ARTHUR ANDERSEN & Co.

Chicago, Illinois,
February 12, 1969.

EXHIBIT 1
AMERICAN ECONOMIC ASSOCIATION
STATEMENT OF ASSETS AND LIABILITIES
DECEMBER 31, 1968

ASSETS		LIABILITIES	
CASH.....	\$ 47,369	ACCOUNTS PAYABLE.....	\$ 49,019
INVESTMENTS (see note):		RESERVE FOR HANDBOOK.....	40,000
- U. S. Government obligations and corporate bonds, at market value (cost of \$124,817).....	\$115,834	DEFERRED INCOME:	
Corporate stocks, at market value (cost of \$289,196).....	515,419	Life membership dues.....	\$ 31,300
		Prepaid membership dues.....	17,570
		Prepaid subscriptions.....	35,214
RECEIVABLES, less allowance for doubtful accounts of \$2,600.....	631,253		84,084
ADVANCES, relating to United States Government grant for foreign student screening and selection.....	30,150	DEFERRED INCREMENT IN MARKET VALUE OF INVESTMENTS (see note).....	94,698
PREPAID EXPENSES.....	4,470	RESTRICTED FUND BALANCES (Exhibit 2).....	36,669
OFFICE FURNITURE AND EQUIPMENT, at cost, less accumulated depreciation of \$3,126.....	1,910	GENERAL FUND:	
	4,067	Balance, December 31, 1967, as previously reported.....	\$433,081
		Less—Adjustment to reflect change in method of accounting for investments (see note).....	(12,269)
			\$420,812
		Balance, December 31, 1967, as adjusted.....	414,749
		Less—Excess of expenses over income for the year ended December 31, 1968 (Exhibit 3).....	(6,063)
			\$719,219

The accompanying note to financial statements is an integral part of this statement.

NOTE TO FINANCIAL STATEMENTS

DECEMBER 31, 1968

Effective December 1, 1966, the Association changed from a cost to a market value basis of accounting for investments. Under the new method, that portion of the increment in market value of corporate stocks which was equivalent to the increase in the underlying retained earnings of those stocks was reflected in income currently and the balance of the increment, after adjusting for an inflation factor, was to be recognized in income over a three-year period.

During 1968, the Association modified the method of recognizing income relating to corporate stocks by eliminating the retained earnings factor from the formula. Accordingly, only the increment in market value of corporate stocks, after adjusting for an inflation factor, is recognized in income over a three-year period. The retroactive effect of this change in method of accounting for investments resulted in reducing the General Fund balance at December 31, 1967, by \$12,269 and increasing the Deferred Increment in Market Value of Investments by \$12,269. The excess of income over expenses for the thirteen-month period ended December 31, 1967, was reduced by \$535.

The change in market value of U. S. government obligations and corporate bonds, after adjusting for an inflation factor, is reflected in income currently.

EXHIBIT 2

STATEMENT OF CHANGES IN RESTRICTED FUND BALANCES FOR THE YEAR ENDED DECEMBER 31, 1968

	Balance Dec. 31, 1967	Re- ceipts	Disburse- ments	Balance Dec. 31, 1968
The Ford Foundation grants for- Translation of foreign economic publications and survey of foreign economic research.....	\$25,045	\$ —	\$(10,058)	\$14,987
Foreign student screening and selection.....	20,187	492	(8,833)	11,846
United States government grant for foreign student screening and selection.....	7,269	—	(2,799)	4,470
The Asia Foundation for Asian economist membership dues to the American Economic Association and related travel expenses.....	2,190	—	(799)	1,391
The National Science Foundation grant for visiting scientist program...	3,157	7,559	(6,968)	3,748
Sundry grants.....	229	—	(2)	227
	<u>\$58,077</u>	<u>\$8,051</u>	<u>\$(29,459)</u>	<u>\$36,669</u>

EXHIBIT 3

AMERICAN ECONOMIC ASSOCIATION STATEMENT OF INCOME AND EXPENSES FOR THE YEAR ENDED DECEMBER 31, 1968

INCOME FROM DUES AND ACTIVITIES:

Membership dues.....	\$160,102
Subscriptions.....	70,561
Advertising.....	48,794
Sale of copies, republications and handbooks.....	5,442
Sale of mailing list.....	12,008
Annual meeting.....	16,993
Sundry.....	6,594

INVESTMENT INCOME (see note):

U. S. government obligations and corporate bonds-			\$320,494
Interest.....	\$ 8,444		
Less- Decline in market value (see note).....	7,804	\$ 640	
Corporate stocks-			
Cash dividends.....	\$11,231		
Increment in market value recognized (see note).....	32,752	43,983	44,623
Total income.....			<u>\$365,117</u>

EXPENSES:

Publication (Exhibit 4).....	\$264,152		
Administrative (Exhibit 4).....	98,628		
Grant to Yale University for <i>Index of Economic Journals</i>	8,400	371,180	
Excess of expenses over income for the year.....			<u>\$ (6,063)</u>

The accompanying note to financial statements and Exhibit 4 are an integral part of this statement.

REPORT OF THE AUDITOR

507

EXHIBIT 4

AMERICAN ECONOMIC ASSOCIATION
STATEMENT OF PUBLICATION AND ADMINISTRATIVE EXPENSES
FOR THE YEAR ENDED DECEMBER 31, 1968

PUBLICATION EXPENSES:

Printing and mailing of-	
<i>American Economic Review</i>	\$112,808
<i>Journal of Economic Abstracts</i>	53,194
<i>Papers and Proceedings</i>	36,637
Provision for <i>Handbook</i>	8,000
Addressing expenses.....	6,676
Editorial salaries.....	29,203
Payments to contributors.....	2,102
Editorial supplies and expense.....	5,208
<i>Journal of Economic Literature</i>	10,324
Total publication expenses.....	<u>\$264,152</u>

ADMINISTRATIVE EXPENSES:

Secretary's salary, less \$3,600 allocated to restricted funds.....	\$ 9,667
Office salaries.....	42,984
Office machine rentals.....	724
Postage.....	3,037
Stationery, printing and supplies.....	3,434
Insurance.....	596
Annual meeting expenses.....	7,635
Executive Committee expenses.....	2,180
Committee on Economic Education expenses.....	758
Other committees' expenses.....	1,002
Archives.....	250
Pension annuity payments.....	7,854
Payroll taxes.....	2,241
Depreciation of furniture and equipment.....	570
Telephone and telegraph.....	1,292
Committee on Research and Publications.....	2,341
President and president-elect expenses.....	2,000
Sundry.....	11,870
Expenses allocated to restricted funds.....	(1,807)
Total administrative expenses.....	<u>\$ 98,628</u>

AMERICAN ECONOMIC REVIEW

REPORT OF THE MANAGING EDITOR FOR THE YEAR ENDING DECEMBER 1969

The number of manuscripts submitted during 1969 was 758, 121 more than the preceding year and more than double the number submitted in 1963. The number submitted has been growing at 18 to 19 percent each year since 1966; and since 1959 at an instantaneous rate of 10 percent. These figures are summarized in Table 1, which shows the number of manuscripts submitted and published since 1948. During the past year 121 papers were published. The sharp jump in number published is due to the change in the *Review's* format, and the transfer of sections on Book Reviews, and Titles of New Books to the *Journal of Economic Literature*.

The distribution of submitted and published manuscripts by subject matter is shown in Table 2. As in previous years the two most popular individual fields are micro-theory and international economics. Almost all fields are represented in the *Review*, and the lack of representation indicates either a low level of interest (history of thought and methodology) or the possibility of classification bias on the editor's part (administration and business finance). There has been a growth in the num-

TABLE 1
MANUSCRIPTS SUBMITTED AND PUBLISHED, 1948-69

Year	Submitted	Published	Ratio of Published to Submitted
1948	205	48	.23
1949	200	53	.27
1950	197	53	.27
1951	222	49	.22
1952	190	47	.25
1953	234	51	.22
1954	231	47	.20
1955	245	41	.17
1956	242	48	.20
1957	215	40	.19
1958	242	46	.19
1959	279	48	.17
1960	276	46	.17
1961	305	47	.15
1962	273	46	.17
1963	329	46	.14
1964	431	67	.16
1965	420	59	.14
1966	451	62	.14
1967	534	94	.18
1968	637	93	.15
1969	758	121	.16

ber of manuscripts submitted in the area of applied welfare economics; that is, in welfare programs, consumer economics, and urban

TABLE 2
SUBJECT-MATTER DISTRIBUTION OF SUBMITTED AND PUBLISHED MANUSCRIPTS IN 1969

	Submitted	Published
General economics—theory; history; systems		
General economics and general equilibrium theory.....	20	3
Microeconomic theory.....	204	28
Macroeconomic theory.....	38	6
Welfare theory.....	78	12
History of thought; methodology.....	3	—
Economic systems.....	6	2
Economic growth; development; planning; fluctuations.....	64	14
Economic statistics.....	16	2
Monetary and financial theory and institutions.....	62	11
Fiscal policy and public finance.....	38	5
International economics.....	88	17
Administration; business finance.....	—	—
Industrial organization.....	10	2
Agriculture; natural resources.....	11	1
Manpower; labor; population.....	43	12
Welfare programs; consumer economics; urban and regional economics.....	61	5
Other.....	16	1

TABLE 3
SUMMARY OF CONTENTS 1968-69

	1968		1969	
	Number	Pages	Number	Pages
Articles.....	30	600	47	579
Communications, including notes, comments, replies....	63	287	67	303
Special articles.....			7	55
Total.....	93		121	
Dissertations.....		63		46
Notes and announcements.....		67		47
Books reviewed.....		424		—
New books.....		82		—
		1,523		1,030

economics. I expect this field to continue to grow and hopefully the proportion of good manuscripts to increase.

Table 3 summarizes the contents of the four regular issues of the *Review* during 1969, and compares it with the previous year. The page counts are not strictly comparable because of the change in format. The economy of the new format is readily seen. The same content occupies fewer pages. The average article occupied 20 pages in 1968 and 12 pages in 1969; the average communication occupied 5 pages in 1968 and the same in 1969. I do not expect the average size of communications to decline in the future, because it is a convenient method of presenting shorter articles. The special

articles shown in Table 3 are seven papers presented at the December, 1969, meetings which were published in the September and December, 1969, issues. The purpose of their publication was to provide readable papers to the membership in advance of the meetings, to stimulate interest and comment. There is some question whether this is a satisfactory method of providing papers in advance, in view of the space taken up and the consequent need to postpone publication of manuscripts which have been refereed and edited.

Table 4 shows actual annual expenditures from 1965 through 1969 for the four regular issues of the *Review*, and the estimated budgets for 1969 and 1970.

TABLE 4
ACTUAL AND BUDGETED EXPENDITURES, 1965-70

	Actual Expenditures					Budget	
	1965	1966	1967	1968	1969	1969	1970
Printing and mailing.....	\$61,606	\$ 74,277	\$ 82,840	\$ 92,948	\$ 97,183	\$ —	\$104,701
Payments to contributors.....	1,598	1,873	2,255	2,140	600	—	—
Subtotal.....	\$63,204	\$ 76,151	\$ 85,095	\$ 95,088	\$ 97,783	\$ —	\$104,701
<i>Office Expenses</i>							
Editor's salary.....	\$ 9,000	\$ 9,000	\$ 9,000	\$ 9,000	\$ 9,000	\$ 9,000	\$ 9,000
Editorial assistance.....	12,024	14,189	16,218	17,772	14,617	14,666	18,339
Assistant to editor.....	6,619	7,330	8,025	8,465	8,416	8,416	10,935
Book reviews.....	2,000	2,000	2,000	2,205	—	—	—
Secretarial.....	1,725	2,160	2,400	2,565	5,038	4,000	5,819
Clerical, proofreading etc....	1,679	2,699	3,793	4,536	1,163	2,250	1,585
Postage and supplies.....	1,462	1,479	1,895	2,382	2,522	2,400	2,977
Office equipment.....	—	175	—	—	626	485	639
Subtotal.....	\$22,486	\$ 24,844	\$ 27,114	\$ 29,155	\$ 26,765	\$26,551	\$ 30,955
Total.....	\$85,691	\$100,996	\$112,209	\$124,243	\$124,548		\$135,656

TABLE 5
COPIES PRINTED, SIZE AND COST OF PRINTING AND MAILING IN 1969

	Printed	Pages		Cost		
		Net	Gross	Issue	Reprints	Total
March.....	27,000	220	304	\$22,384.53	\$203.23	\$22,587.77
June.....	27,000	236	268	20,733.19	138.44	20,871.63
September.....	27,000	292	336	26,065.79	187.06	26,252.85
December.....	27,000	282	368	27,250.00*	221.00*	27,471.00*
	108,000	1,030	1,276	\$96,433.51	\$749.73	\$97,183.25

* Estimate.

With the change in format of the *Review* we did not estimate a printing budget for 1969. For the same reason it is not possible to provide a standard for judging the minor increase of \$4,235 in printing expenses for 1969 over 1968. The total (gross) number of pages printed was 1,276 in 1969 and 1,828 in 1968; the change due to the new two-column format. The reduced number of pages has led to lower press and material costs. But the increased number of manuscripts printed has led to increased typesetting charges. We printed 121 manuscripts in 1969 and 93 in 1968. Printing costs will rise again in 1970 due to a rise in basic charges made by the printer and to an expected increase in number of copies printed. I expect printing and mailing expenses to total \$104,701 in 1970. There is no budgeted item for payments to contributors, as this practice has been discontinued.

Budgeted office expenditures for 1969 were \$26,551, while actual expenditures were \$26,765. The 1970 budget is shown in the last column of Table 4. Office expenses will increase to \$30,955, mainly due to salary increases and increased proofreading assistance. The increase for postage and supplies is made in anticipation of continued growth of manuscript submissions.

Table 5 gives information on printing and mailing costs by issues during 1969. There were 27,000 copies per issue printed in 1969, compared to 26,744 in 1968 and 12,400 in 1958. Increases are of course due to the greater number of members and subscribers.

Two members of the Board of Editors complete their three-year terms at this time: Dale

Jorgensen and Ronald W. Jones. I wish to thank them personally for their fine cooperation, hard work, and high standards of scholarship. The entire Association owes them a debt of gratitude. In addition, Herbert Stein, who was nominated to the Board December, 1969, found he was unable to serve due to the pressure of his duties at the Council of Economic Advisers.

Last year the size of the Board was increased from ten to twelve members. In anticipation of increasing the number to fourteen, I would like to nominate the following individuals for the terms shown:

Warren Smith, to December 31, 1970—one year

Marc Nerlove, to December 31, 1972—three years

Gregory Chow, to December 31, 1972—three years

Jagdish Bhagwati, to December 31, 1973—four years

Barbara Bergmann, to December 31, 1973—four years

I should like to express my thanks to the members of the Board of Editors whose terms are continuing: Gary Becker, Carl Christ, Charles Ferguson, Irwin Friend, Harry Johnson, Alvin Marty, Warren Nutter, Vernon Smith, and Lloyd Ulman. I have received valuable proofreading and mathematical assistance from Martin Ringo, Bijan Aghevli, and William Holahan. Finally, I should like to thank Mrs. Wilma St. John for her fine work as assistant editor.

In addition to the members of the Board, I

have sought the assistance of a large number of economists, during the course of the year. I wish to thank them for their patience and cooperation in reading and evaluating manu-

scripts. They have provided advice and talents not represented on the Board and have eased their heavy work load. The following have assisted in this way:

H. Aaron	G. Brown	T. Finn	D. Kamerschen
I. Adelman	J. Brown	A. Fisher	M. Kamien
M. Adelman	M. Brown	J. Floyd	N. Kaplan
D. Aigner	H. Bruton	D. Foley	D. Katzner
G. Akerlof	J. Buchanan	A. Friedlander	R. Kavesh
P. Albin	E. Burmeister	G. Fromm	D. Keesing
L. DeAlessi	J. Burns	V. Fuchs	M. Klein
R. Aliber	P. Cagan	D. Gale	A. Klevorick
P. Allen	G. Cain	L. Gallaway	A. Kneese
W. Allen	C. Campbell	R. Gemmill	L. Krause
E. Altman	W. Carleton	N. Georgescu-Roegen	M. Kreinin
S. Altman	W. Chang	C. Goetz	A. Krueger
E. Ames	H. Chenery	S. Goldfeld	E. Kuh
J. Anderson	V. Chetty	S. Goldman	K. Kurihara
W. Anderson	G. Childs	M. Gordon	P. Kuznets
S. Arndt	J. Chipman	R. Gordon	R. Lampman
K. Arrow	B. Chiswick	M. Gort	K. Lancaster
O. Ashenfelter	G. Chow	H. Grabowski	W. Landes
H. Averch	L. Christensen	D. Granick	M. Landsberger
W. Baer	R. Coase	H. Grossman	L. Lave
M. Bailey	A. Conrad	M. Grove	T. Lee
J. Bain	R. Coen	J. Guttentag	N. Leff
B. Balassa	R. Cooper	J. Hadar	R. Leftwich
R. Baldwin	P. Cootner	H. Hagemann	A. Leijonhufvud
R. Barlow	J. Cragg	R. Hall	A. Lerner
R. Barro	J. Cross	R. Hamada	E. Lerner
Y. Barzel	M. Daniels	M. Hamburger	J. Letiche
W. Baumol	E. Davis	W. Hansen	C. Leven
M. Beckmann	O. Davis	J. Hanson	D. Levhari
J. Behrman	F. de Leeuw	R. Haveman	H. Levin
B. Bergmann	E. Denison	A. Heins	F. Levy
J. Berliner	R. Dernberger	J. Henderson	S. Lewis
S. Besen	D. Dewey	B. Hickman	H. Liebfafsky
J. Bhagwati	J. Dirlam	J. Hirshleifer	T. Liu
R. Bodkin	A. Dixit	H. Hochman	M. Lovell
V. Bonomo	E. Domar	C. Holt	I. Lowry
S. Bowles	R. Dorfman	G. Horwich	R. Lucas
J. Bowman	D. Dugan	F. Hsiao	R. McKean
M. Bowman	D. Durand	S. Hymans	L. McKenzie
W. Brainard	R. Dusansky	M. Intrilligator	J. McKie
H. Brems	G. Eads	D. Jaffee	R. McKinnon
J. Brittain	R. Eckaus	G. Johnson	B. Malkiel
R. Britto	E. Fama	L. Johnson	A. Manne
M. Bronfenbrenner	P. Feldstein	M. Johnson	E. Mansfield
E. Brown	W. Fellner	T. Juster	J. Margolis

T. Marschak	D. North	R. Ruffin	L. Thurow
T. Mayer	H. Nourse	V. Ruttan	T. Tideman
D. Meiselman	E. Olsen	H. Ryder	R. Timberlake, Jr.
A. Meltzer	D. Orr	A. Sametz	J. Tobin
J. Melvin	L. Orr	R. Sato	M. Todaro
C. Meyer	R. Parks	J. Schaefer	A. Treadway
P. Meyer	A. Pascal	M. Schupack	E. Truman
R. Meyer	D. Patinkin	T. Schultz	S. Tsiang
C. Michalopoulos	M. Pauly	C. Schultze	G. Tullock
M. Miller	S. Peltzman	D. Schwartzman	J. Ullman
N. Miller	R. Penner	D. Schydrowsky	D. Usher
J. Milliman	G. Perry	T. Scitovsky	J. Vanek
E. Mills	C. Phelps	G. Scully	W. Vickrey
J. Mincer	E. Phelps	L. Shapley	T. Wales
H. Minsky	G. Pierson	W. Sharpe	F. Walker
E. Mishan	C. Plott	W. Shepherd	R. Waud
F. Modigliani	R. Portes	M. Shubik	R. Weil
M. Montias	R. Powell	B. Siegel	P. Weinstein
J. Morgan	F. Puffer	W. Silber	R. Weintraub
L. Moses	J. Quirk	L. Sjaastad	B. Weisbrod
J. Mossin	L. Rapping	E. Smolensky	L. Weiss
D. Mueller	A. Rees	J. Stein	M. Weitzman
R. Muth	S. Reiter	P. Steiner	F. Welch
M. Myers	S. Resnick	G. Stigler	F. Westfield
M. Nadiri	R. Rhomberg	J. Stiglitz	T. Whitin
K. Nagatani	G. Rich	M. Straszheim	M. Whitman
S. Naya	A. Rivlin	A. Strout	O. Williamson
P. Neher	A. Robichek	R. Summers	R. Wilson
R. Nelson	E. Rolph	A. Takayama	J. Wood
M. Nerlove	H. Rose	A. Tang	C. Wright
P. Newman	S. Rosen	L. Taylor	M. Yaari
D. Nichols	J. Rosse	L. Telser	P. Zarembka
J. Niehans	J. Rothenberg	A. Throop	

GEORGE H. BORTS, *Managing Editor*

JOURNAL OF ECONOMIC LITERATURE

REPORT OF THE MANAGING EDITOR FOR THE YEAR ENDING DECEMBER 1969

Although the Executive Committee received and approved my submission for 1970 at its March, 1969, meeting, I am taking advantage of a point I mentioned then that after the *J.E.L.* office had been functioning a year, I would be in a better position both to discuss the *Journal's* contextual form and its budgetary requirements. Hence, this report goes beyond the March statement on contents and it contains a revised budget request.

Volume VII, *Journal of Economic Literature*

Contents. The 1969 volume of *J.E.L.* is numbered Volume VII in order, for reasons of postal permits, to keep the authorization given by the Evanston, Illinois, postmaster to the *Journal of Economic Abstracts*.

Volume VII contained 1,432 pages. There were four general survey articles, 2 "creative curmudgeon" articles expressing clearly a normative point of view, 5 review articles, several communications, "normal" length reviews of 188 books, 1,207 annotations of books, tables of contents listings from 668 journal issues (we list over 200 journals), classifications of about 4,200 articles (with about 6,000 total listings), and about 1,100 abstracts of articles.

The number of journals which we will list fully or partially has been increasing, both as a result of our contacting journals on the list (as inherited from Professor Smithies) but not submitting material and requests for inclusion from new journals or journals only partially devoted to economic topics.

Costs. Circulation of the issues in this volume averaged about 25,000.

An inventory of about 3,000 copies per issue was printed at the time of appearance of each issue. These copies will be eventually sold at \$2.00 per issue. Insofar as I am aware, the printer (Banta) stores this inventory at no cost to the Association. Mailing cost is about 14 cents per issue. Thus, the value of the inventory is about \$24,000 less expected mailing costs of \$1,700 leaving an inventory with a future net value of \$22,300.

One principal problem during the first year of any journal is production schedule. Because of the size of *J.E.L.* and particularly the uncertainties associated with number of words (and spaces) and type sizes, our anticipated gross space requirements came to about 150 more pages than I had originally anticipated. It would have been more had we not slightly reduced type size in some of the departments. The greater size plus the problems of learning the operation also caused delays both in the Editorial Office and at the printer. The sorry record is that the March issue came out in June; the June issue in September; and the September issue in early October; and the December issue will come out in early January, 1970.

Printer's and mailing costs for the four issues came to about \$97,000. Assuming that the inventory (less potential mailing costs) amounts to \$22,300, it is appropriate to conclude that the real direct cost for printing and distribution to regular subscribers came to about \$74,700.

Editorial costs are not easy to allocate. I write as of December 19, 1969, a time when five issues of *J.E.L.* have gone to press but when only four have been proofread and generally processed. We have also commissioned (and paid advances on one) additional survey article. Editorial costs of *J.E.L.* preparation contain much that will appear in Volume VIII of the *Index of Economic Journals*. And, finally, I, the editor, underestimated the amount of editorial time required in a bibliographical journal of the scope and size of *J.E.L.* and I arranged, because of a special relationship with the Associate Editor (working at pay assumed to cover 17½ hours per week) to jolly her into working in excess of 40 hours per week with promises of better things to come. The exploitation of the Assistant Editor was less marked, but still significant. In all events, editorial costs for calendar 1969 were (as compared to budgeted amounts):

	Expended*	Budgeted
Editor.....	\$ 9,000	\$ 9,000
Editorial staff.....	27,700	26,500
Associate editor		
Assistant editor		
Secretarial work		
Annotation, classification and additional assistance		
Supplies and communications	3,500†	2,700
Commissions for articles....	3,500‡	3,000
	\$43,700	\$41,200

* Does not include fringe benefits.

† Includes \$425 future posting.

‡ Includes partial payment for 1970.

Future Contents

1. Survey articles have been "commissioned" on the following topics: distribution of income, J. Mincer; radical economics, M. Bronfenbrenner; productivity and production functions, H. I. Nadiri; agricultural economics, B. F. Johnston.

2. Discussions which might result in "creative curmudgeon" articles are in process with: Warren Smith, Martin Shubik, Franco Modigliani, Harry Johnson, Robert Eisner. Others have been asked, but as these are "contributed" pieces, receipt is uncertain.

3. Length of abstracts has been reduced to 100 words. Attempts have been made to try to increase the appearance of abstracts from rel-

atively less accessible journals.

4. I anticipate that Volume VIII (1970) will have about 1,424 pages, including the annual index.

5. According to the printer's estimates, costs of printing 28,000 copies each (and mailing about 25,000) of the four issues of the *J.E.L.* in 1970 will be about \$104,000.

BUDGET FOR 1970

	1969 Expended	1970
Managing Editor (part-time)...	\$ 9,000	\$ 9,000
Associate Editor (full-time)....	5,500*	15,000
Assistant Editor (full-time)....	8,100	10,000
Secretarial Service (2 full-time)...	8,500	11,000
Annotation, classification and editorial assistance.....	5,600	6,000
Supplies and communication....	3,500	2,600
Commissioned articles.....	3,500	2,500
	\$43,700	\$56,100

* Paid on part-time basis.

Editorial Board

During the year Professor Marc Nerlove replaced Professor Henry Houthakker on the Editorial Board, joining Professors Richard A. Easterlin, David I. Fand, Karl A. Fox, Arnold C. Harberger, and Peter B. Kenen.

MARK PERLMAN, *Managing Editor*

REPORT OF THE COMMITTEE ON ECONOMIC EDUCATION

During 1969, this Committee has continued to encourage, in all the ways it is able, further attention to the improvement of teaching at the undergraduate and secondary school levels. As in the past, we have worked closely with the Joint Council on Economic Education, which, as an operating agency, has cooperated in implementing a number of the recommendations made by this Committee. Major activities for the year include:

1. A new *Journal of Economic Education* has been established, as the first journal in the area of economic education. It is published by the Joint Council on Economic Education, with Professor Henry Villard of this Committee as editor and Dr. John Maher of the J.C.E.E. as managing editor. This Committee serves on the editorial board, and the policies of the *Journal* have been developed cooperatively between this Committee and the J.C.E.E. The first edition appeared in October. We are pleased to report that there is already a paid subscription list approaching 2,000; and the supply of manuscripts for consideration has far exceeded our expectations. We view this *Journal* as a major contribution to the exchange of information on teaching experiments and research in the general area.

2. On June 10-15, a special interdisciplinary institute (the "Grove Park Institute") was held for leaders in the various social science disciplines to consult on the entire problem of education for future teachers in the social studies. Bernard Haley (Stanford University) served as chairman of the economics group; others were Kenneth Boulding, Ben Lewis, Henry Villard, and Harold Williamson from the A.E.A. Several recommendations of the Institute are included in the following items.

3. At the suggestion of the Grove Park Institute and this Committee, the Joint Council on Economic Education is mailing out to each department of economics in the country a report and recommendations to the California State Department of Education on "College Preparation for Teaching Economics." This report was developed by a group including members of this Committee and other educa-

tors at the university and secondary school levels. It is concerned especially with the training of future teachers of economics in the primary and secondary schools.

4. A special experiment to test the proficiency in economics of a sample of current graduates of teacher-training institutions has been prepared and submitted by the University of Wisconsin-Milwaukee to the Office of Education, following up a recommendation of the Grove Park Institute and this Committee. Henry Villard serves as the A.E.A. representative on the policy board overseeing this group of projects.

5. The Committee voted to support the recommendation of the Grove Park Institute that universities throughout the country consider carefully the possibilities of offering a special M.A. degree in economics for future teachers of economics at the secondary school level, presumably in cooperation with departments or schools of education on their campuses. We also supported the Grove Park recommendation to the Office of Education that special funding be considered for such programs and for teachers in them.

6. Two "pilot plant" projects in the teaching of elementary economics at the college level are in process, with reports to be submitted by Professor Rendigs Fels at Vanderbilt and Professor Phillip Saunders at Carnegie-Mellon. These reflect the Committee's recommendation that special experimental courses be devised, with complete reporting and evaluation, which may be of assistance to departments which do not have the resources to develop such courses on their own.

7. Following the development of the Test of Economic Understanding (TUCE) last year, the J.C.E.E. in cooperation with this Committee is establishing a national "question bank." Faculty members throughout the country will be invited to send test questions to the question bank where they will be screened, classified by type, and kept on hand to meet requests from economics teachers throughout the country. Only questions which have been tested carefully by the originating institution will be put in the question bank.

To encourage the development of questions designed especially to test problem-solving ability in economics and the application of economic principles to complex problems, special prizes (reaching \$1,000 for first prize) will be awarded next year for those questions which prove most imaginative and successful within this general category. Financing for the contest is being provided by Vanderbilt University out of a J.C.E.E. grant for the pilot plant course in economics there. The J.C.E.E. will handle the administration of the contest.

8. The Committee has under consideration the establishment, next year or the following year, of a special summer workshop in the teaching of economics for outstanding third- and fourth-year graduate students in economics who are going out to teaching posts without substantial experience. The goal is to help a group of some 30-40 leading young economists get started well in teaching, where under present arrangements they often get little assistance; and to encourage them to use planned experimentation and careful evaluation processes for courses that they may introduce at the various schools to which they go.

9. The Visiting Scientists Program, financed by the National Science Foundation, has continued to operate effectively in 1969. A special report by Professor Phillip Saunders, who is the administrator of the program under the general policy guidance of this Committee, is attached.

10. In the programming of the sessions on economic education at the annual meetings of the Association, this Committee has continued to press toward papers reporting careful research work in the field and especially interesting projects rather than mere hortatory statements to do better. We are encouraged that the number of strong candidates for inclusion on these sessions has risen steadily over the past several years.

11. In November, the American Economic Association was presented a special citation by the Joint Council on Economic Education on the occasion of the J.C.E.E.'s twentieth anniversary. The A.E.A. has worked closely with the Joint Council throughout this period, especially during the past decade, and it received the citation for outstanding contributions to the J.C.E.E.'s programs and to the improvement of economics teaching on all fronts. The citation was accepted by this Committee on behalf of the Association. The relationship with the J.C.E.E. has been a fruitful one from the Association's viewpoint. Our working relationships have been close, and without the cooperation of the J.C.E.E. it would have been impossible to carry out the wide-ranging list of activities which this Committee has undertaken over the past decade, toward improving the teaching of economics and teachers in economics at both the primary-secondary school and college levels.

G. L. BACH, *Chairman*

REPORT ON THE VISITING SCIENTISTS PROGRAM IN ECONOMICS

Academic year 1968-69 marked the third year of the American Economic Association's participation in the Visiting Scientists Program sponsored by the National Science Foundation. The purposes of the program are: (1) to stimulate interest in economics among undergraduates in colleges and universities through visits by leaders in the profession;

tee on Economic Education, Professor G. L. Bach, of Stanford University, Chairman. All financial transactions were executed and recorded through the offices of the A.E.A.'s Secretary-Treasurer, Professor Harold Williamson.

The following summary figures cover the first three years of the program:

School Year	Number of Schools Visited	Number of Visit Days	Expenditure	Cash Contributions	Cost to NSF
1966-67	20	34	\$4,903	\$223	\$4,680
1967-68	32	43	5,274	684	4,590
1968-69	29	49½	5,596	990	4,606

and (2) to provide opportunities for economics staff members and administrators at the visited institutions to discuss their teaching and research problems with the visiting economist. Emphasis is placed on visits to smaller colleges and universities whose major focus in economics is at the undergraduate level.

As in the preceding year, the 1968-69 program was administered by Professor Phillip Saunders, of Carnegie-Mellon University, under the direction of the Association's Commit-

The 1968-69 visits covered 16 different states from Massachusetts to California and from New York to Texas. In arranging visits to the schools selected, the informal roster of distinguished economists established in the preceding year was used extensively, and it was supplemented with direct arrangements in a few specific cases. The continuing cooperation and enthusiasm of the individual economists asked to make visits is one of the most heartening aspects of the program.

PHILLIP SAUNDERS

REPORT OF THE JOINT COMMITTEE OF THE AMERICAN ECONOMIC
ASSOCIATION AND THE ASSOCIATION OF AMERICAN
LAW SCHOOLS

The Joint Committee on Law and Economics sponsored by the A.E.A. and the A.A.L.S. had an eventful year, highlighted by a "research consultation meeting" held on March 18, 1969. The research consultation meeting was sponsored by the Walter E. Meyer Research Institute and was held at Stanford University. The subject was, "Products Liability: Trends and Implications." Twenty-seven professors of economics and professors of law attended this meeting. A major monograph was written for this meeting by Ronald McKean. Written critiques of this monograph were also prepared and delivered by Professors James Buchanan, Robert Dorfman, Guido Calabresi, and Grant Gilmore. The Committee intends to publish this monograph and the critiques. At present, discussions are being held with possible publishers.

The Committee has also engaged in a continuing discussion of various activities which may further the joint exploration of legal-economic problems. Consideration is now being given to another research consultation meeting on the subject of "Property Rights in Non-Patentable and Non-Copyrightable Ideas."

There is also a desire to go beyond this sponsorship of annual meetings which bring together professors of law and economics on a

subject of common interest. The hope is to have an eventual impact on some forms of curriculum revision in schools of law and possibly also economics departments. To this end we not only favor publication of the papers from the annual meetings, but also a special meeting of economists who teach in law schools. These economists may well profit from sharing their common problems and noting some opportunities. Institutional support may eventually be requested for the advancement of teaching and research in this area.

The Walter E. Meyer Research Institute of Law has declared a moratorium on research support while the Institute undergoes reorganization. As a result, the Committee is in the process of seeking support from other sources. The future activities of the Committee are now therefore dependent on obtaining this additional support. Thus, it is respectfully requested by your Committee that permission be given by the Executive Committee of the American Economic Association and the Executive Committee of the Association of American Law Schools to proceed with efforts to gain financial support for the activities of this Committee.

GERALD M. MEIER, *Chairman*

REPORT OF THE COMMITTEE ON THE FUNCTIONS OF THE ANNUAL MEETING

We feel, after some reflection and discussion, that the annual meeting does perform certain functions and cannot be easily dispensed with. This is partly evidenced by the preferences revealed through attendance, even by those who are neither speakers nor on the labor market. There seem to be four functions served by the meeting: (1) labor market for entrants into the profession; (2) job mobility for those who have been already entered; (3) informal communication among members of the profession, both intellectually and socially; and (4) formal intellectual communication through the program.

The increasing size of the meetings has led to deterioration in the performance of these functions. All communication has become increasingly difficult and costly, in particular hotel message systems do not have the capacity to handle high propensity message-sending demands when they are fully booked. Informal communication also becomes increasingly chancy. It is clear, however, that nothing desirable can be done about reducing the size of the meetings, since any step in this direction would detract from their value.

There seems to be a widespread consensus that the formal programs are relatively unsuccessful. The attempt to follow a specific theme has not really worked out in practice. A large proportion of the papers are executed as a matter of routine and obligation rather than out of a desire to present new viewpoints to the economic world.

Recommendation 1. The meetings should be primarily devoted to the presentation of con-

tributions to be selected from among those submitted by a program committee, to be appointed by the President-elect. Emphasis should be given to younger members of the profession to give them maximum possibility for mobility. However, some senior men will doubtless be needed to attract more listeners. In addition to the Presidential Address and the Ely Lecture, the President-elect should organize two or three sessions of invited papers. Possibly, some provision should be given for the circulation of abstracts of the contributed papers, while invited papers could still be published in the *Review*.

Recommendation 2. The physical organization of the entering labor market could be improved. The Association should make itself responsible for creating a communication system between eager graduate students and economics department recruiters. A desk, separate from the hotel communication system, should perhaps be set up especially for this purpose. Possibly all recruiters should be given rooms together or at least in the same hotel. It may be possible to assess costs against the departments concerned.

It probably follows from the above that the *Papers and Proceedings* become superfluous. It is not clear with the flood of communications that there is any real value to an unrefereed journal. If it is thought desirable, the funds thus saved might be transferred to increasing the number of issues in the *American Economic Review*.

KENNETH J. ARROW, *Chairman*

REPORT OF THE REPRESENTATIVE TO THE SOCIAL SCIENCE RESEARCH COUNCIL

Dr. Henry W. Riecken was elected President of the Social Science Research Council on January 1, 1969, succeeding Pendleton Herring, who had served in this position for twenty years. Mr. Herring has become director of the Foreign Area Fellowship Program.

The Council joined with other research organizations in trying to bring about modification of proposed legislation on the taxation of foundations so that it will not adversely affect the work of nonprofit scientific research organizations and learned societies.

The report of the Behavioral and Social Sciences Survey Committee, jointly sponsored by S.S.R.C. and the National Academy of Sciences—National Research Council, was published in October, 1969. Carl Kaysen and Robert M. Solow served as members of the central planning committee for the survey and as chairman and co-chairman, respectively, of its panel on economics. The report of this panel will be distributed early in 1970.

Under the S.S.R.C.'s Committee on Economic Stability, chaired by Bert Hickman, work was begun on Project LINK. This project will attempt to forge links, representing

international trade and financial flows, between econometric models of national economies. The Committee on the Economy of Mainland China has completed its program because further work was made impracticable by difficulties of access to Chinese economic data.

A Conference on Research on Urban Poverty, jointly sponsored by the S.S.R.C. and Princeton University, was held in Princeton in May, 1969. The conference provided an occasion for the exchange of views among economists, sociologists, and political scientists engaged in research on problems of the urban poor.

The Council's programs of pre- and post-doctoral research training fellowships and faculty research grants are continuing to attract many well-qualified applicants. The fellowship program now emphasizes support for candidates with interdisciplinary interests. The Council is also administering a small program of grants to support research by social scientists who are members of ethnic minority groups or on the problems of such groups.

ALBERT REES

REPORT OF THE REPRESENTATIVE TO THE NATIONAL BUREAU OF ECONOMIC RESEARCH

The National Bureau of Economic Research was organized in 1920 in response to a growing demand for objective determination of the facts bearing upon economic problems and for their interpretation in an impartial manner. It is a private, nonprofit organization. Its Board of Directors includes men of different economic and social viewpoints with members from industry and labor, banking and finance, professional associations of economists, accountants and statisticians, and economists from university faculties. The American Economic Association is one of the representative organizations which nominate a member of the National Bureau's Board of Directors.

The Bureau is supported by grants, often for specific research projects, from foundations, trade associations, and government agencies, by annual contributions and subscriptions, and by interest and dividends on sustaining funds. Its expenditures in the last fiscal year, ending June 30, 1969, exceeded \$2 million for the first time.

Staff and Conference Programs. During 1969 staff research at the National Bureau of Economic Research continued in the following general areas: economic growth—productivity, employment, price behavior, interrelated factor demand functions, and tax structures; national income, consumption and capital formation; urban and regional economic studies; human resources and social institutions; business cycles, short-term forecasting and econometric model simulations; financial institutions and processes; studies in industrial organization—including the service industries, the health-care and medical industries, and studies of managerial incomes and stockholder returns; international studies—including the relation of U.S. manufacturing abroad to U.S. exports, the role of prices in international trade and balance-of-payments adjustment policies.

Among new studies undertaken were those on the benefits of higher education by Paul Taubman and Terence Wales; law enforcement and the courts by William M. Landes;

an updating and deepening of flow-of-funds data and national balance sheets by Raymond W. Goldsmith; the use of credit in Communist countries by George Garvy; econometrics and measurement methods by Yoel Haitovsky and Thomas J. Sargent.

The National Bureau is continuing to improve its data sources and computer usage. Its time series data banks are being computerized to facilitate dissemination to interested parties. A new group, an Inter-institutional Conference on the Application of the Computer to Economic Research, is being organized to deal with computer problems. It is expected that this group, operating along the lines of the Conference on Research in Income and Wealth, will meet periodically to consider common problems, exchange information, and produce conference reports on developments that are occurring in this field.

The Conference on Research in Income and Wealth sponsored a meeting on Econometric Models of Cyclical Behavior at Harvard University on November 14–15, 1969. A Conference on International Comparison of Prices and Real Income is planned for the spring of 1970.

The Universities-National Bureau Committee for Economic Research has arranged a Conference on the International Mobility and Movement of Capital to be held in early 1970. A Conference on Education as an Industry is planned for the fall of 1970.

As part of the commemoration of the National Bureau's fiftieth anniversary in 1970, the Board and staff are planning an Anniversary Dinner in the spring and a series of colloquia to be held during the fall of 1970. Six areas of economic research are expected to be assessed with respect to past efforts and future significance in the light of policy issues: economic growth, business cycles, finance and money markets, industrial organization, human resources, and public expenditures and taxation. A conference is planned for late 1970, including the colloquia participants, Bureau staff, and others, with the objective of synthesizing the views developed in the collo-

quia and evaluating future research priorities.

Publications. Eight National Bureau publications were issued in 1969. The following were in press at the end of 1969: *The Design of Economic Accounts*, by Nancy and Richard Ruggles; *Home Mortgage Delinquency and Foreclosure*, by John P. Herzog and James S. Earley; *Monetary Statistics of the United States: Estimates, Sources, Methods*, by Milton Friedman and Anna J. Schwartz; *Errors in Provisional Estimates of Gross National Product*, by Rosanne Cole; *The Technology Factor in International Trade*, Report of Universities-National Bureau Committee for Economic Research, Raymond Vernon, Editor; *Education, Income, and Human Capital*, Report of Conference on Research in Income and Wealth, W. Lee Hansen, Editor; *The Behavior of Industrial Prices*, by George J. Stigler and James K. Kindahl; "The Seasonal Variation of Interest Rates," by Stanley Diller; "The Changing Position of Philanthropy in the American Economy, 1929-1959," by Frank G. Dickinson; "Economic Factors in the Growth of Corporation Giving," by Ralph L. Nelson; "The Value of Time in Passenger Transportation: The Demand for Air Travel," by Reuben Gronau.

Interim reports on the current activities of the Board, staff, conferences, new studies, available manuscripts and reprints, were presented in the form of a newsletter, *National Bureau Report*, which was published in June and December, 1969. These reports were accompanied by supplements which dealt at greater length with items of current interest, viz., "The New ASA-NBER Survey of Forecasts by Economic Statisticians," by Victor Zarnowitz, and a statement by John R. Meyer which was presented before the Special Subcommittee on the Evaluation and Planning of Social Programs of the Senate Committee for

Labor and Public Welfare.

Meetings of the Board of Directors in 1969. A special meeting of the Board was held on May 9, 1969, to review developments in the research program. Particular attention was given to proposals for new research efforts that were presented in three reports: Analyzing the Effects of Large-Scale Changes in Fiscal Structure: A Proposed Systems Approach, by John Bossons and Carl S. Shoup; International Studies: Changing Patterns of International Comparative Advantage, by Hal B. Lary; and Opportunities for Research on National Balance Sheets, Flow-of-Funds Models, and Other Financial Topics, by Raymond W. Goldsmith.

At the Annual Meeting, September 19-20, 1969, the bylaws were amended so that the fifteen Directors on the Bureau's Board who are appointed by universities could be chosen from among a larger number of universities than heretofore. Previously the universities' representative Directors were nominated by each of the same fifteen universities that had been designated by the Board. Under the new procedure, the Bureau's nominating committee may invite nominations from any of the thirty-seven universities that is a member of the Universities-National Bureau Committee for Economic Research. The Board also designated the Canadian Economics Association as an organization entitled to nominate a Director.

The National Bureau's Forty-Ninth Annual Report, *New Challenges for Economic Research*, was presented to the Board at the Annual Meeting. Copies of the report are available upon request from the National Bureau of Economic Research, 261 Madison Avenue, New York, New York 10016.

WILLARD L. THORP

REPORT OF REPRESENTATIVE TO THE INTERNATIONAL ECONOMIC ASSOCIATION

During 1969 two round table conferences were held: one in Kandy, Ceylon, June 2-10, on "Economic Development in South Asia," the other in Algarve, Portugal, August 28 to September 4, on "The Mutual Repercussions of North American and Western European Economic Policies." Members of the American Economic Association who participated were Jagdish N. Bhagwati, Benjamin I. Cohen, and Mark W. Leiserson at the conference in Ceylon, and Richard N. Cooper, John H. Dunning, Charles P. Kindleberger, Robert A. Mundell, and Edmund S. Phelps at the conference in Portugal.

Three conferences and a training workshop are planned for 1970: on "The Essence of a Growth Model" (Israel, March); "Market Relations and Planning" (Prague, Czechoslo-

vakia, May); "The Increasing Gap in the Income per Capita between the Rich and the Poor Nations" (open, August); and the "First European Training Workshop" for some twenty-five young economists (Louvain, Belgium, August-September).

Five volumes were published in 1969 in English and two in French. The English ones are: *Problems of Agriculture in Industrial Societies and Repercussions in Developing Countries*, edited by G. U. Papi and C. C. Nunn; *Risk and Uncertainty*, edited by K. Borch and J. Mossin; *The Analysis of the Public Economy*, edited by J. Margolis; *Backward Areas in Advanced Countries*, edited by E. A. G. Robinson; and *The Future of International Economic Relations*, edited by P. A. Samuelson.

Fritz Machlup

REPORT OF THE ECONOMICS INSTITUTE POLICY AND ADVISORY BOARD

During the year 1969 the Economics Institute has successfully continued its valuable service of providing a transitional summer program for foreign students who have been admitted to graduate study in economics and agricultural economics in the United States. One hundred and ten students attended this year's session. They came from 32 countries and held admission to 35 different U.S. institutions of higher learning. The Economics Institute has now completed twelve sessions. During that time a total of 894 students representing 78 countries have participated in the program prior to undertaking graduate work in 104 universities in the United States.

The financial condition of the Economics Institute continues to be sound. Its underwriting revolving fund—invested by the American Economic Association—stands at about \$160,000. Tuition and fee receipts in 1969 totaled about \$97,000 as compared with \$80,000 in 1968.

Two new members—Professors Gustav Ranis, of Yale, and Jeffrey Williamson, of Wisconsin—joined the Policy and Advisory Board at its Boulder meeting on October 6 and 7. Since the training program of the Institute was now running smoothly and efficiently, the director and the board members concentrated on two longer-range issues: the Institute's role in continuing the services previously offered by the A.E.A. screening project and the ways in which the Institute experience might be brought to bear on efforts being made to expand higher educational opportunities for members of minority groups in this country. The screening project will be carried on by the Institute on the modified and experimental basis for a few years. The much larger problem of educational opportunities for minority groups is now under intensive study by a subcommittee and we hope to report on this matter in detail next year.

HENRY ROSOVSKY, *Chairman*

REPORT OF THE ASSOCIATION'S REPRESENTATIVES TO THE GROVE PARK INSTITUTE

The Grove Park Institute, sponsored by CONPASS, was held in Asheville, North Carolina, June 10-15, 1969, and was attended by representatives of various learned societies and educational associations. The Committee representing the American Economic Association at the Institute included the following: Bernard F. Haley, Chairman, Kenneth E. Boulding, Moe L. Frankel, Ben Lewis, Henry Villard, Arthur Welsh, and Harold F. Williamson; Committee Secretary, Gertrude Tait.

Based on the discussions at the Institute of the problems of improving the quality of teachers and the teaching of economics, the Committee made the following recommendations:

1. That CONPASS or the Office of Education fund a test of proficiency in economics of a selected sample of current graduates of teacher training institutions who have been prepared to teach courses with an economics content at the secondary level.

2. That the American Economic Association give its support to the implementation of the Report and Recommendations of an Advisory Seminar to the California State Department of Education (*College Preparation for Teaching Economics*) which sets forth appropriate course requirements to be fulfilled in the preparation for teaching economics and social studies at the precollege level. If it is impractical to require specific courses, we recommend that comparable proficiency be determined by acceptable standardized tests.

3. That the American Economic Association consider affiliation with the Association of Organizations for Teacher Education. In addition we recommend that whatever agency undertakes the project of testing the proficiency in economics of a selected sample of current graduates of teacher training institutions should work with the A.O.T.E. or other appropriate organizations in the implementation of this project.

4. That departments of economics recognize their obligations to provide one or more economics courses at the senior or graduate

level designed to meet the needs of prospective teachers of social studies at the secondary level.

5. That a pilot project be undertaken to test the feasibility of developing an integrated social studies course at the junior or senior level for prospective social studies teachers. It is proposed that CONPASS finance or seek financing for the project. In addition, CONPASS is asked to invite the American Association of Colleges of Teacher Education to appoint one or more consultants to assist the proposed committee.

6. That there be continued provision of an M.A. program in economics (or economics and economic education) for prospective teachers of economics in the secondary schools and junior colleges. We urge the Office of Education to provide more support than has recently been made available to economics M.A. candidates.

7. That consultant services to social studies teachers at both the secondary and junior college level be continued and strengthened. Academic institutions and economics departments should encourage their faculty members to serve as consultants to social studies faculties at the elementary, secondary, and junior college levels.

8. That funding be made available to construct a "National Question Bank."

9. That the American Economic Association undertake a review of the undergraduate curriculum in economics with a view to advising its members in colleges and universities with respect to desirable changes to meet current criticisms of present-day students and faculty that in our judgment may well have merit.

10. That there be developed an alliance and dialogue between those members of departments of economics interested in the improvement of teaching and appropriate members of departments or schools of education who have something to contribute on methodology and learning theory. This might well be facilitated if a joint appointment could be

made in each institution of a professor of economics and education. This faculty member would not only serve as liaison but might also assume an active role in the department of economics in heading up a program of teacher training for graduate students.

11. That graduate schools require all students for the Ph.D. degree to have at least the

equivalent of two quarters of teaching experience before receiving the degree. This experience should be under the supervision of a senior faculty member, and such methods as the use of tape recordings of the student's performance in class, followed by critical analysis by the supervisor, should be employed.

HENRY H. VILLARD

REPORT OF THE INFORMAL MEETING OF THE BLACK CAUCUS AND AMERICAN ECONOMIC ASSOCIATION REPRESENTATIVES

An informal meeting of members of the Black Caucus and American Economic Association representatives, attended by approximately sixty-five individuals, was held at the New York Hilton Hotel on the evening of December 28, 1969. Discussion was centered on the items listed in copies of a "Suggested Bill of Particulars" prepared by the Black Caucus and circulated at the meeting. In brief, the Bill of Particulars called on the A.E.A. to: (1) appoint an Associate Secretary, designated by the Black Caucus, whose major function would be to give direction to activities designed to improve the status of economists from black and other ethnic groups; (2) implement a plan to increase the number of black economists, including a program to provide financial assistance; (3) arrange for greater participation by black economists in the policy-making and programming activities of the Association; (4) provide financial support for program to be administered by the

Black Caucus designed to increase the number of undergraduates majoring in economics; and (5) to publish a "Statement of Concern" prepared by the Black Caucus in the 1970 *Papers and Proceedings* (see below).

Following the discussion, A.E.A. President Fellner announced that the Executive Committee had already authorized the appointment of a committee, including representatives of the Black Caucus, to consider plans for increasing the supply of black economists. He also indicated that further consideration would be given to the other items in the Bill of Particulars at the next meeting of the Executive Committee.¹

HAROLD F. WILLIAMSON
Secretary Pro Tempore

¹ At its meeting on March 6-7, 1970, the Executive Committee voted to assist the Caucus of Black Economists in organizing and obtaining funds for a program to promote professional research activity by minority-group economists related to "Economic Development of Ethnic Communities."

STATEMENT OF CONCERN OF THE
CAUCUS OF BLACK ECONOMISTS TO THE
AMERICAN ECONOMIC ASSOCIATION*

In recognition of the gross underrepresentation of black economists in the life of the Association and the profession, the members of the Caucus express with urgency this statement of concern. The grievances of blacks and other minorities, who have been restricted in their professional development by the practices of the Association as well as those of the institutions which it serves, form the main part of our concern. What the Association has done or has failed to do has had negative implications and severe adverse effects on minority economists and minority communities as a whole. Their consequences give rise to our forthright criticism and challenge.

In response to this situation substantial changes in policies, practices, and procedures are in order. Some of the needed changes have already been suggested to the officers of the Association.¹ This statement consists of the central issues pertaining to our concern. Further elaboration of the issues and proposals for change will be provided through the activities, position papers, and other work of the Caucus of Black Economists.

Social and Racial Bias. Social and racial bias can be seen in many covert and overt forms in Association policies and practices. Most obvious is the insensitivity and indifference to the position of blacks and other minorities in the profession. In its racial dimensions the disproportionate representation of blacks and other minorities in the profession has been taken to mean an absence of ability to compete and achieve rather than the presence of discrimination which restricts the op-

portunity to compete and achieve. Such an attitude, held even by economists in honored positions in the profession, has served to rationalize and perpetuate the inequities resulting from discrimination based on social and racial bias. Until the real causes and cures for the situation are recognized and appropriate changes made, the professional development of blacks will continue to be blocked, significant talent will go unutilized, and many economists from minority groups will remain underemployed.

Professional Bias. Although vast socioeconomic differences persist between whites and nonwhites and the latter are plagued with immobilities of various kinds, the economics profession has not given adequate recognition to the theoretical and policy implications of these differences. While individual economists have documented the fundamental error of applying the economics of affluence to areas that call for an economics of poverty, the Association has not taken cognizance of and acted upon these differences.

Form and technique have been favored over substance and values while issues of policy, perspective, and institutional performance have been disregarded or downgraded in professional terms. In its apolitical zeal, the Association has remained silent on the crucial questions and decisions regarding resource utilization affecting minorities. Further, professionalism and academic detachment have been used as shields for barring meaningful responses to urgent social problems. Thus, the Association has been professionalized into inaction regarding the fundamental problems of resource utilization and wealth distribution, especially as regards minority communities. Yet gross disparity and discrimination along racial lines are familiar conditions throughout the profession as well as in the economy.

In view of the injustice which the traditional posture of the Association condones, a reinterpretation of the role of economist as professional is long overdue. By design and default the economic well-being of large num-

* Written in behalf of the Caucus of Black Economists. As a member of the Steering Committee of the Caucus and Association Chairman for 1970, the author takes full responsibility for its contents. This procedure was agreed upon at the organizing meeting of the Caucus in December, 1969. Members of the Steering Committee are: Marcus Alexis, Richard F. America, Jr., Robert D. Bell, Robert S. Browne, Horace R. Carby-Samuels, Cleveland A. Chandler, Flournoy Coles, Frank G. Davis, Karl D. Gregory, Thaddeus H. Sprattlen, Robert C. Vowels, and Charles Z. Wilson.

¹ Presented in the Bill of Particulars of the Caucus, December 29, 1969.

bers of the minority population is being sacrificed for the illusion of neutrality on social and political questions. Clearly, past experience suggests that "professional neutrality" perpetuates the *status quo*. By practicing "neutrality" the power acquired through professional knowledge and its uses is placed primarily at the disposition of the privileged classes and their institutions.

Lack of Social Responsibility. Even though the utilization of scarce resources constitutes the core of the discipline, the uses of the knowledge and power of economics has not been dealt with by the Association. Its social purposes have not been equated with its social power. This constitutes a serious lack of social responsibility on the part of the profession as a whole. Even the immediate question of racial discrimination against economists has received little attention by the Association. Yet it is employment discrimination that prevents significant numbers of blacks from access to high-yield professional positions. Hence, the Association has not dealt with the missing link in parity competition between whites and nonwhites in the profession. Of course this represents an extension of the restrictive practices that still are apparent in doctoral programs, the top of the hierarchy in professional education.

The Challenge. The Association is now being called upon to face up to its responsibilities as the representative body of a profession. It is being challenged to correct its errors of omission and commission in its policies and practices. Since the officers and other members of the Association are fully aware of the importance of resources in accelerating growth and change, the imperative response called for is the allocation of resources sufficient to correct the inequities that are perpetuated through the failures and errors of the economics profession. To do so, the priorities of its representative organization will have to change. So will the perception of the key issues—social, racial, and professional—confronting the Association.

If the challenge is met, social and racial bias can be contained. With sufficient resolve and resources devoted to the task, professional bias can be eliminated and a full measure of social responsibility achieved.

The Caucus of Black Economists is ready and willing to assist the Association in redirecting its efforts and priorities as well as in implementing programs that meet the needs of blacks and other minority groups in the profession and the economy.

THADDEUS H. SPRATLEN, *Associate Chairman*

PUBLICATIONS

OF THE

AMERICAN ECONOMIC ASSOCIATION

1970

AMERICAN ECONOMIC REVIEW

Address all inquiries regarding publications of the Association from 1886 through 1910 and Vols. I-LIII (1911-63) of the *American Economic Review* to: Walter J. Johnson, Inc., 111 Fifth Avenue, New York, New York 10003. Inquiries regarding subsequent issues of the *Review* should be addressed to the American Economic Association office. The supply of those items double starred is exhausted.

The titles of papers published in the *Papers and Proceedings*, as well as the titles of articles in the *American Economic Review*, are indexed in the *Reader's Guide* to 1953 and the *International Index* thereafter.

Volume LIV, 1964

The American Economic Review, **March, **June, **September, and **December;	
each	2.00
Supplement (No. 1—January).—Handbook. Pp. 464.†	8.00
**Supplement (No. 3—May).—Seventy-sixth Annual Meeting:	
Papers and Proceedings: Richard T. Ely Lecture; Theory of Monopolistic Competition after Thirty Years; Principles of Efficiency; Efficiency in Agriculture; Efficiency in the Labor Markets; Regulated Industries; Efficiency in the Use of Natural Resources; Efficiency in the Government Sector; Problems in Taxation; Financial Factors in Business Cycles; Reappraisals in American Economic History; Comparative Costs and Economic Development; Canadian-American Economic Relations—Efficiency and Independence; Efficiency of the Soviet Economy; Invited Dissertations, I and II; Efficiency in the Teaching of Economics. Pp. 698.	4.00
Supplement (March).—Surveys of Foreign Postwar Developments in Economic Thought. Pp. 55.	1.00
Supplement (June).—Surveys of Foreign Postwar Developments in Economic Thought. Pp. 101.	1.00

† This issue is not included in the price of the volume.

Volume LV, 1965

The American Economic Review, **March, June, September, and December; each \$2.00
Supplement (No. 2—May).—Seventy-seventh Annual Meeting:

Papers and Proceedings: Richard T. Ely Lecture; Economic Growth—International Comparisons; Comparative Economic Systems—Nationalized Industry; Economic History—Its Contribution to Economic Education, Research, and Policy; Contribution of the History of Economic Thought to the Understanding of Economic Theory, Economic History, and the History of Economic Policy; Evolving International Monetary Mechanism; Domestic Implications of the Evolving International Monetary Mechanism; Survey Research—Three Surveys, Findings and Implications for Theory and Policy; Recent Capital and Production Theory; Technological Change—Stimuli, Constraints, Returns; New National Econometric Model—Its Application; Invited Student Dissertations; Defense Economics—Applying Economic Criteria; Nonmarket Decision Making; Economic Theory and Nonprofit Enterprise; Economics of Poverty; Economic Education—Experiments in the Teaching of Economics. Pp. 652.

4.00

Supplement (December).—The Structure of Economists' Employment and Salaries, 1964. Pp. 98.

1.00

Volume LVI, 1966

(\$10.00 a Volume)

The American Economic Review, March, June, September, and December; each 2.00
Supplement (No. 2—May).—Seventy-eighth Annual Meeting:

Papers and Proceedings: Richard T. Ely Lecture; Allocation and Distribution Theory—Technological Innovation and Progress; Capital Theory—Technical Progress and Capital Structure; Economic Development—Advanced Technology for Poor Countries; Knowledge, Information, and Innovation in the Soviet Economy; Money and Banking—Innovations in Finance; Public Finance—Promotion of Knowledge Production and Innovation; International Economics—Progress and Transfer of Technical Knowledge; Antitrust and Patent Laws—Effects on Innovation; Public Regulation—The Impact of Changing Technology; Economics of Education; Economics of Publishing; Economics of Broadcasting and Advertising; Economics of Science Policy; Production and Use of Economic Knowledge; Labor Economics—Effects of More Knowledge. Pp. 670.

4.00

Supplement (September).—Handbook. Pp. 184.

2.50

Volume LVII, 1967

The American Economic Review, March, June, September, and December; each 2.00
Supplement (No. 2—May).—Seventy-ninth Annual Meeting:

Papers and Proceedings: Richard T. Ely Lecture; Military Manpower Procurement; Cost-Benefit Analysis for Government Decisions; Economics of Health; Economic Analysis of Water Resource Problems; Transportation and Patterns of Urban Development; Antitrust and Monopoly; Economic History; Economic Development; Property Rights and Behavior; Topics in Money; Aggregate Models; Measurement of Price Change; Invited Doctoral Dissertations; *Das Kapital*—A Centenary Appreciation; Efficiency of Education in Economics. Pp. 741.

4.00

Volume LVIII, 1968

The American Economic Review, March, June, September, and December; each 2.00
Supplement (No. 2—May).—Eightieth Annual Meeting:

Papers and Proceedings: Richard T. Ely Lecture; Behavioral and Ecological Economics; The Entrepreneur; Econometric Models, Their Problems and Usefulness; Regional Economic Models; Invited Student Dissertations; Problems in Pricing and Growth; Nonmarket Decision Making; Transport for Economic and Social Development; Economics of Arms Control and Disarmament; Experiments in Teaching Economics; Technological and Economic Implications of 3 Percent Growth; Economics and Noneconomics of Poverty; Economic Reform in Eastern Europe and the U.S.S.R.; International Liquidity. Pp. 751.

4.00

Supplement (June).—Surveys of National Economic Policy Issues and Policy Research. Pp. 138. 1.00

Supplement (December).—Studies of the Structure of Economists' Salaries and Income. Pp. 153. 1.00

Volume LIX, 1969

The American Economic Review, March, June, September, and December; each Supplement (No. 2—May).—Eighty-first Annual Meeting: 2.00

Papers and Proceedings: Richard T. Ely Lecture; Theory of Innovation; Research and Development and Other Determinants of Investment; Theory of the Firm and of Market Structures; Wage-Price Dynamics, Inflation, and Unemployment; Supply Function in Agriculture Revisited; Problems in the Theory of Public Choice; Efficiency of Education in Economics; Transportation and the Public Utilities; Monetary Theory; Future of Gold; Round Table on Exchange Rate Policy; Economic History; Economic Development; Problems in the Area of Poverty; Strategic Theory and Its Applications; Round Table on Allocation of Resources in Law Enforcement; Centralization and Decentralization in Economic Systems; Completed Dissertation Research. Pp. 629.

4.00

Supplement (September).—Surveys of National Economic Policy Issues and Policy Research. Pp. 118. 1.00

Supplement (No. 6).—Handbook. Pp. 603.† 10.00

† This issue is not included in the price of this volume.

Volume LX, 1970

The American Economic Review, March, June, September, and December; each Supplement (No. 2—May).—Eighty-second Annual Meeting: 2.00

Papers and Proceedings: Richard T. Ely Lecture; Investment Behavior; Money within the General Framework of the Economic System; Programming Models of National Economics; Industrial Organization: Retrospect and Prospect; New Dimensions in Natural Resources; Basic Data for Policy and Public Decisions: Technical Aspects; Planning to Free Enterprise Economies; National Communications Policy; Economics of Imperialism; Distribution Issues: Trends and Policies; Dynamics of Income Distribution: Poverty and Progress; Market and Price Mechanism in Socialist Countries; Economics of Education; Teaching of Economics; Increasing the Supply of Black Economists; International Firm and Efficient Economic Allocation; New Ideas in Pure Theory; Economy of Cities; Economics in the Industrial State: Science and Sedative. Pp. 533.

4.00

JOURNAL OF ECONOMIC LITERATURE

Since 1966, membership in the Association has included, along with the *American Economic Review*, a subscription to the *Journal of Economic Abstracts*, now the *Journal of Economic Literature*. Vols. I-V (1963-67) of the *Journal of Economic Abstracts* may be ordered from Walter J. Johnson, Inc., 111 Fifth Avenue, New York, New York, 10003. Vol. VI (1968) of the *Journal of Economic Abstracts* and Vol. VII (1969) of the *Journal of Economic Literature* may be ordered through the American Economic Association office.

* * *

The American Economic Association, founded, among other purposes, for the "encouragement of economic research" and "the encouragement of perfect freedom of economic discussion," has over nineteen thousand members, including public and professional men and most of the leading students of political economy in America. Membership dues are ten dollars a year. Each member receives all current reports and publications of the Association, including the *Journal of Economic Literature*. In addition to members, there are some seven thousand subscribers to the *American Economic Review*, institutional and individual. Applications for membership and other inquiries should be sent to:

Harold F. Williams

SECRETARY OF THE AMERICAN ECONOMIC ASSOCIATION

629 Noyes Street, Evanston, Illinois 60201

Economic Texts

from the college department
Rand McNally & Company



New 1970

INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE

BY F. M. SCHERER, *UNIVERSITY OF MICHIGAN*

A new conceptual approach to industrial organization analysis emerges from this valuable new text. Integrating structural and behavioral models, the book explores problems economists have not yet solved.

Fully accessible to students with a basic grasp of price theory, the book is specifically concerned with variables affecting the quality performance of industrial firms in a market economy: resource allocation and use, distributive equity, progressiveness, and macroeconomic stability.

Outstanding as a text for advanced industrial organization and public policy, this book is addressed to other audiences as well — policy-makers, businessmen, and professional economists.
c. 535 pages Order Number 67-6509 \$13.00

Widely Adopted

INTRODUCTION TO MODERN MICROECONOMICS

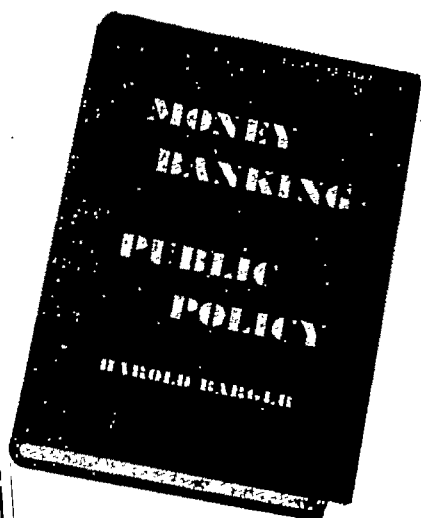
BY KELVIN LANCASTER, *COLUMBIA UNIVERSITY*

With this widely used text, Professor Lancaster puts the science of modern microeconomics squarely into the undergraduate curriculum. This text presents an elementary account of modern microeconomics. Students have nothing to unlearn when they take advanced courses; the book concentrates on principles rather than method. Mathematics is kept simple.
326 pages Order Number 67-6251 \$8.95

New 1970

STUDENT WORKBOOK FOR INTRODUCTION TO MODERN MICROECONOMICS

By David L. Schulze, *Iowa State University*
c. 208 pages Paper Order Number 67-6252 \$3.95



Second Edition

MONEY, BANKING AND PUBLIC POLICY

BY HAROLD BARGER, *COLUMBIA UNIVERSITY*

All the topics traditional to a college course on money and banking, and several uncommon ones, make Professor Barger's text preeminent in this field. The new edition, completely revised and updated, includes a chapter on international monetary reform and expanding treatment of monetary equilibrium. 760 pages. Order Number 67-6247 \$9.95
Instructors Manual available.

MONEY AND THE BALANCE OF PAYMENTS

BY TIBOR SCITOVSKY, *YALE UNIVERSITY*

The demand and supply of money—stated in terms of modern consumer demand theory—and the theory of payments adjustment—the first to appear in book form—are the major topics explored in this recently published text. Written for students with a general knowledge of economics, it is simply argued and easy to approach.

188 pages. Order Number 67-6250 \$8.95

THE DEMAND AND SUPPLY OF PUBLIC GOODS

BY JAMES M. BUCHANAN

VIRGINIA POLYTECHNIC INSTITUTE

Here is the first systematic, full-length treatment of the modern theory of public goods. The models presented are specifically related to the theory of collective decision-making.

214 pages. Order Number 67-6707 \$8.95

Examination Copies: If you wish to consider adoption of Rand McNally college texts, please write on departmental stationery indicating your academic position, the course title, approximate enrollment, and the text currently in use.

COLLEGE DEPARTMENT

Rand McNally & Company

P.O. Box 7600 Chicago, Illinois 60680

*A completely new approach in which
theory and practice are integrated and made truly
applicable to the student's world*

PRINCIPLES OF ECONOMICS

By **Daniel B. Snits**
The University of Michigan

Offering a new approach to introductory economics, that is less mathematical than those of most principles texts, this book has been six years in development and has been thoroughly class-tested. The determinants of productive capacity are examined via the live setting of world poverty, and demographic and economic factors are integrated into a unified dynamic analysis. Special chapters deal with poverty in the U.S. and with the economics of environmental pollution. By studying the actual behavior of the economy, students are led to do their own forecasting. Monetary and fiscal policy are integrated into a unified theory of stabilization policy that shows the two as complements rather than rivals. One hundred years of U.S. economic history are surveyed as a case study of economic development. *March, 1970. 543 pp.; \$9.95*

To accompany the text: Study Guide (\$2.95) and Instructor's Manual.



1817

Send for our 1970 catalogue/#36-00475

HARPER & ROW, Publishers,
49 E. 33d Street, New York 10016

ANNOUNCING

GRADUATE STUDIES IN ECONOMICS

Revised Edition, 1969
\$2.50

The American Economic Association has recently published a revised edition of the **GRADUATE STUDIES IN ECONOMICS: A Guide to Graduate Programs in Economics and Agricultural Economics at United States and Canadian Universities**. This volume is designed to provide prospective graduate students and their advisors with up-to-date information on requirements, admission procedures, programs, costs, fellowships, etc., of departments in the United States and Canada offering programs for advanced degrees. Reports on 150 graduate departments are included in the revised edition of the *Guide* plus a brief description of economics as a field for graduate study.

Order from

Richard D. Irwin, Inc.
1818 Ridge Road
Homewood, Illinois 60430

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Van Tassel Reorganizes Economic Principles Course

Students at Clark University are learning economic principles by means of "deferred payoff." They spend the first semester of the introductory course learning the theory that is usually taught in an entire year's course, then they enjoy a second semester of discussion about policy and events.

Roger Van Tassel, Chairman of the Department of Economics at Clark University, has been teaching the introductory course this way for several years. He uses his own textbook: **ECONOMIC ESSENTIALS: A Core Approach**.

"We hit them rather hard with economic theory and not much of anything else in the first semester," he said in a recent interview. "Then we give them a fair exposure to real world problems, and they can deal with these problems better having mastered a certain amount of economic theory."

"You no longer have to begin by showing them that economics is all around them," he says. "Students now *know* that economics is important. Likewise, it makes more sense to take them through all the theory you want them to learn before giving them any policy analysis."

Currently about one fourth of students entering Clark University have studied economics in high school, and Professor Van Tassel predicts that in another four or five years the majority of freshmen will have studied some economics before entering college.

Uses Core Text

ECONOMIC ESSENTIALS: A Core Approach, by Professor Van Tassel, is the student's primary reading in the first semester of the introductory course. A 410-page paperback, it was published by Houghton Mifflin in the spring of 1969.



JULIE O'NEIL

Roger Van Tassel, Chairman of the Department of Economics at Clark University, Worcester, Mass., discussed his reorganized principles course during recent interview.

The book deals almost exclusively with economic theory and its applicability to an understanding of economic policy and events. It omits historical and institutional material and detailed accounts of policy.

"In this area," Professor Van Tassel says, "there are short books and there are good rigorous

books but I don't think there are any other rigorous, short books."

"You can go into the year, or in my case, the second semester," Professor Van Tassel says, "not really knowing what you're going to talk about in terms of the policy events and leave half the coverage open on the basis of what the devil seems to be going on."

ECONOMIC ESSENTIALS: A Core Approach

Roger C. Van Tassel, Clark University

410 pages / paper / 1969 / \$4.95

**Houghton
Mifflin**

Boston / Atlanta / Dallas / Geneva, Ill. / New York / Palo Alto

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

Since publication in February, 1969, more than 140 colleges and universities have adopted the Second Edition of

ECONOMICS

By Richard G. Lipsey, University of Essex

Peter O. Steiner, The University of Michigan

Lipsey and Steiner's **ECONOMICS** has, in its two editions, reflected the profound change in method and point of view that the discipline has undergone during the past two decades. The approach, the method, and the structure of this book have been planned to teach economics as it is practiced—as the genuine social science it has become. The format of the **SECOND EDITION** is entirely new. It includes expanded treatment of *Macroeconomics* (rewritten, reorganized, and more than doubled in size, with the number and variety of examples substantially increased); *Policy* (major policy discussions and materials expanded and better integrated); *Demand Theory* (now includes marginal utility theory, indifference theory, and revealed preference theory); *Public Policy Toward Monopoly* (with a discussion of public utility regulation); *Fiscal Policy*; and *Growth Underdeveloped Economies* (rewritten to incorporate much up-to-date comparative data). There are many new chapters, including *Price Theory in Action*, *Monetary Policy*, and *Comparative Economic Systems*. 845 pp.; \$10.95

To accompany the text: Study Guide by Dorothy Forbush and Dascomb R. Forbush (341 pp.; \$3.50) and Instructor's Manual by Lipsey and Steiner.



1817

Send for our 1970 catalogue/#36-00475

HARPER & ROW, Publishers,
49 E. 33d Street, New York 10016

1969 HANDBOOK of the AMERICAN ECONOMIC ASSOCIATION

The 1969 HANDBOOK of the American Economic Association contains biographical data on approximately 13,000 members, a geographical listing of members, and listings of members by fields of specialization and academic affiliation. The volume is \$10.00 a copy, postage prepaid. Orders should be sent to:

AMERICAN ECONOMIC ASSOCIATION

629 Noyes Street

Evanston, Illinois 60201



european community

COMMON MARKET

*Publications and
Statistics Available*

EUROSPECTRA

The scientific and technical quarterly (formerly *Euratom Review*) of the European Communities. Contains articles on European developments in biology, agricultural and medical uses of nuclear energy, chemistry, health protection, industrial applications of radioisotopes and radiation, mathematics and computers, metallurgy, physics, reactor technology, waste treatment and storage, legal and economic aspects of atomic energy, and automatic documentation. Published quarterly in English, German, French, Italian, Dutch. *Annual subscription \$3.50*

BULLETIN OF THE EUROPEAN COMMUNITIES

Official monthly bulletin containing news of the activities of the European Communities Commission and other institutions of the three Communities: Economic (EEC), Coal and Steel (ECSC), and Atomic Energy (Euratom). Provides coverage of intra- and extra-Community affairs, and includes texts of important memoranda or programs in frequent supplements to the Bulletin. In English, German, French, Italian, Dutch. *Annual subscription (Bulletin, Supplement, Index) \$5.00*

GRAPHS AND NOTES ON THE ECONOMIC SITUATION IN THE COMMUNITY

Monthly report on the Communities' economic situation, covering industrial production, unemployment, consumer prices, and the

balance of trade. Each quarter, the report includes data on other subjects, such as gold and foreign exchange reserves, wages, prices, trade, sales, tax revenue. In three bilingual editions: English/French, German/Dutch, and French/Italian. *Annual subscription \$6.00*

THE ECONOMIC SITUATION IN THE COMMUNITY

A quarterly survey by the European Communities Commission analyzing the Communities' economic situation, demand, production and balance of markets, and economic data for each of the Community countries. Included is the Commission's readjusted quarterly and annual economic forecasts and observations on short-term economic policy. Complementing the text are graphs and tables on Community trade and economic indicators. In English, German, French, Italian, Dutch. *Annual subscription \$8.00*

GENERAL STATISTICS

Monthly bulletin of general statistics contains the principal data on short-term economic trends in the European Community, by subject and by country. The subject section gives data available for the different Community countries and, in a few important cases, for the United States and the United Kingdom. Subjects cover population and labor, national accounts, agricultural and industrial production, internal and foreign trade, transport, prices, wages, and finance. The country tables cover the Community members, and there are tables relating to the United States, the United Kingdom, and countries associated with the Community. *Annual subscription \$11.00*

AGRICULTURAL STATISTICS

Each issue contains statistics and notes on a different aspect of agriculture, ranging from balance sheets on individual products and supply statistics, to analyses of farm labor, agricultural costs and prices, regional statistics, and land use. *Eight to ten issues yearly. Annual subscription \$9.00*

Subscriptions may be ordered through:

The European Community Information Service
Suite 707, 2100 M Street, N.W.
Washington, D.C. 20037.

or sent directly to:

Sales Office for Official Publications
of the European Communities
37, rue Glesener, Luxembourg

All subscriptions cover the calendar year; rates for airmail delivery available on request. Make check payable to the office through which ordered.

Please mention THE AMERICAN ECONOMIC REVIEW When Writing to Advertisers

NOW AVAILABLE!

The latest reports on the history and present status of Negro employment in U.S. industry from the Industrial Research Unit of the Wharton School, University of Pennsylvania.

NEGRO EMPLOYMENT IN BASIC INDUSTRY

A Study of Racial Policies in Six Industries

By Herbert R. Northrup, et al.

An analysis of current racial policies and Negro employment in automobile, aerospace, steel, petroleum, rubber tire, and chemical companies, most of which ranked among the top 50 in the Fortune directory. Outlines current hiring and promotion trends, probing the determinants of Negro employment practices throughout the industry and the factors affecting company policies.

\$10.00

THE NEGRO IN THE PAPER INDUSTRY

By Herbert R. Northrup

Probes the governmental attack on the paper industry's seniority system and the application of the "rightful place" doctrine by courts in the Crown Zellerbach cases.

\$8.50

Other books in the Racial Policies of American Industry series:

THE NEGRO IN THE BANKING INDUSTRY, by Armand J. Thieblot, Jr. **\$5.95**

THE NEGRO IN THE MEAT INDUSTRY, by Walter A. Fogel **\$4.50**

THE NEGRO IN THE TOBACCO INDUSTRY, by Herbert R. Northrup **\$4.50**

THE NEGRO IN THE INSURANCE INDUSTRY, by Linda P. Fletcher **\$5.95**

THE NEGRO IN THE PUBLIC UTILITIES INDUSTRIES, by Bernard E. Anderson **\$5.95**

Available Fall 1970:

THE NEGRO IN THE BITUMINOUS COAL MINING INDUSTRY, by Darold T. Barnum

THE NEGRO IN THE MOTOR TRUCKING INDUSTRY, by Richard A. Leone

THE NEGRO IN THE RAILROAD INDUSTRY, by Howard W. Risher, Jr.

THE NEGRO IN THE SHIPBUILDING INDUSTRY, by Lester Rubin

THE NEGRO IN THE URBAN TRANSIT INDUSTRY, by Philip W. Jeffress

THE NEGRO IN THE LUMBER INDUSTRY, by John C. Howard

THE NEGRO IN THE TEXTILE INDUSTRY, by Richard L. Rowan

THE NEGRO IN THE DRUG MANUFACTURING INDUSTRY, by F. Marlon Fletcher



UNIVERSITY OF PENNSYLVANIA PRESS

3933 Walnut St., Philadelphia, Pa. 19104